PB87-113791

Precision and Reliability of Laboratory Permeability Measurements

Cincinnati Univ., OH

Prepared for

Environmental Protection Agency, Cincinnati, OH

Oct 86

U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA

(Please read Instructions on the reverse before completing)

1. REPORT NO.
EPA/600/2-86/097

2.

3. RECIPIENT'S ACCESSION NO.
PB87 113791/AS

4. TITLE AND SUBTITLE
PRECISION AND RELIABILITY OF LABORATORY PERMEABILITY MEASUREMENTS

5. REPORT DATE
October 1936

6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)
John Bryant, Ph.D., and Andrew Bodocsi, Ph.D.

8. PERFORMING ORGANIZATION REPORT NO.
68-03-3210-03

9. PERFORMING ORGANIZATION NAME AND ADDRESS
University of Cincinnati
Department of Civil and Environmental Engineering
Cincinnati, Ohio 45221

10. PROGRAM ELEMENT NO.

11. CONTRACT/GRANT NO.
68-03-3210-03

12. SPONSORING AGENCY NAME AND ADDRESS
Hazardous Waste Engineering Research Laboratory
Office of Research and Development
U.S. Environmental Protection Agency
Cincinnati, Ohio 45268

13. TYPE OF REPORT AND PERIOD COVERED
Final, 4/6/84-5/20/85

14. SPONSORING AGENCY CODE
EPA/600/12

15. SUPPLEMENTARY NOTES

16. ABSTRACT

A limited set of laboratory test data on clay liner permeabilities was gathered from six sources to create a data bank suitable for a preliminary statistical analysis. The collected data were also used to survey the most commonly used permeameters and testing methods for clay liner permeabilities.

This report briefly describes the methods used to collect and organize the data. The main body of the report discusses the findings from the analysis of the data. Consideration is given to the degree of variability found in replicated permeability tests, the question of sample equilibration with water and chemicals, and the effect of the magnitude of gradient on permeability. Permeability test results using flexible-wall and fixed-wall permeameters are compared. Consideration is also given to the effects of parameters of sample preparation, such as the molding water content, which may be strongly related to permeability, and the amount of variability of permeabilities that may be found in location-to-location sampling from within a source of liner material. In addition, a discussion on statistical considerations in the design and analysis of permeability experiments is presented.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT
RELEASE TO PUBLIC

19. SECURITY CLASS *(This Report)*
UNCLASSIFIED

20. SECURITY CLASS *(This page)*
UNCLASSIFIED

21. NO. OF PAGES
190

22. PRICE

EPA/600/2-86/097
October 1986

PB87-113791

PRECISION AND RELIABILITY OF LABORATORY PERMEABILITY MEASUREMENTS

by

John Bryant
Department of Quantitative Analysis
and Information Systems
University of Cincinnati
Cincinnati, Ohio 45221

and

Andrew Bodocsi
Department of Civil and Environmental Engineering
University of Cincinnati
Cincinnati, Ohio 45221

EPA Contract No. 68-03-3210-03

Project Officer

Jonathan G. Herrmann
Land Pollution Control Division
Hazardous Waste Engineering Research Laboratory
Cincinnati, Ohio 45268

HAZARDOUS WASTE ENGINEERING RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
CINCINNATI, OHIO 45268

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL
INFORMATION SERVICE
SPRINGFIELD, VA. 22161

DISCLAIMER

The information in this document has been funded wholly or in part by the United States Environmental Protection Agency under Contract No. 68-03-3210-03 to the University of Cincinnati. It has been subject to the Agency's peer and administrative review, and it has been approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

FOREWORD

Todays's rapidly developing and changing technologies and industrial products and practices frequently carry with them the increased generation of solid and hazardous wastes. These materials, if improperly dealt with, can threaten both public health and the environment. Abandonded waste sites and accidental releases of toxic and hazardous substances to the environment also have important environmental and public health implications. The Hazardous Waste Engineering Research Laboratory assists in providing an authoritative and defensible engineering basis for assessing and solving these problems. Its products support the policies, programs and regulations of the Environmental Protection Agency, the permitting and other responsibilities of State and local governments and the needs of both large and small businesses in handling their wastes responsibly and economically.

This document describes the variability observed in a collected set of laboratory permeability measurments on clay liner materials and identifies factors which lead to test variability. The intended audience for this document includes researchers, practitioners, and permit reviewers who deal with the testing, design, or analysis of clay liners. For further information, please contact the Land Pollution Control Division of the Hazardous Waste Engineering Research Laboratory.

Thomas R. Hauser, Director
Hazardous Waste Engineering Research Laboratory

ABSTRACT

A limited set of laboratory test data on clay liner permeabilities was gathered from six sources to create a data bank suitable for a preliminary statistical analysis. The collected data were also used to survey the most commonly used permeameters and testing methods for clay liner permeabilities.

This report briefly describes the methods used to collect and organize the data. The main body of the report discusses the findings from the analysis of the data. Consideration is given to the degree of variability found in replicated permeability tests, the question of sample equilibration with water and chemicals, and the effect of the magnitude of gradient on permeability. Permeability test results using flexible-wall and fixed-wall permeameters are compared. Consideration is also given to the effects of parameters of sample preparation, such as the molding water content, which may be strongly related to permeability, and the amount of variability of permeabilities that may be found in location-to-location sampling from within a source of liner material. In addition, a discussion on statistical considerations in the design and analysis of permeability experiments is presented.

This report was submitted in fulfillment of Contract No. 68-03-3210-03 by the University of Cincinnati, Department of Civil and Environmental Engineering, under the sponsorship of the U.S. Environmental Protection Agency. This report covers a period from April 1, 1984, to September 30, 1985.

TABLE OF CONTENTS

Foreword iii
Abstract iv
Figures vii
Tables. x
Acknowledgments. xi

1.0 Introduction 1
2.0 Conclusions and Recommendations 3
3.0 Methods and Procedures for Data Collection and Analysis . . . 6
  3.1 Collection and Organization of Data 6
  3.2 Test Methods and Equipment Used by Data Sources . . 8
  3.3 Data Base Limitations 11
  3.4 Computer Hardware and Statistical Analysis Packages. 13
4.0 Data Analysis Results and Discussion 14
  4.1 Scales of Measurement and Variability. 14
  4.2 Location-to-Location Variability 30
  4.3 Effect of Molding Water Content on Permeability . . 35
  4.4 Equilibration of Water Permeability Tests 42
  4.5 Effect of Gradient 66
  4.6 Comparison of Flexible-Wall and Rigid-Wall Permeamtrs. 82
  4.7 Comments on Equilibration with Chemicals in Rigid-Wall Tests Permeameters 103
5.0 Experimental Plans for Permeability Studies 126
  5.1 Introduction 126
  5.2 Examples and Comparisons of Experimental Plans. . . 126
    5.2.1 References. 126
    5.2.2 Terminology 127
    5.2.3 Completely Randomized Design. 129
    5.2.4 Randomized Complete Block Design . . . 130
    5.2.5 Factorial Experiments 131
    5.2.6 Determining Required No. of Replicates . 132
    5.2.7 Incomplete Block Designs Which Counfound Factorial Effects 133
    5.2.8 Split-Plot Designs 136

References 139

Appendices

A. Tabulation of Log$_{10}$k vs. k 143
B. Statistical Methods for the Analysis of Experimental Data . . . 145
  B.1 Introduction 145

TABLE OF CONTENTS

B.2 Analysis of Variance - Completely Randomized Design . 145
B.3 Analysis of Variance - Randomized Complete Block Design 146
B.4 Analysis of Variance - Two Factor Treatment Structure, Completely Randomized Design 147
B.5 Analysis of Variance - Two Factor Treatment Structure, Randomized Complete Block Design . . . 150
B.6 Determining the Required Number of Replicates - Specific Methods and Examples 150
B.7 Diagnostic Checks and Remedial Measures 153
B.8 Analysis of Variance - Incomplete Block Design with Partial Confounding of Factorial Effects 157
B.9 Analysis of Variance - Split Plot Designs 159
B.10 Permeability Tests of Proposed Borrow Material . . 163
C. Methods to Confirm Equilibrium of Research Permeability Tests . 167

FIGURES

Number Page

1 Permeability test specifications sheet 9
2 Clay type sheet. 10
3 Permeabilities of four clays with .01N $CaSO_4$ water 15
4 Log$_{10}$ permeabilities of four clays with .01N $CaSO_4$ water . . 18
5 Effect of consolidation on the permeability of Sodium Illite Clay. 20
6 Permeability vs. void ratio of Faceville Clay at 321.74 hrs . 21
7 Permeability of Kaolinite Clay with heptane 24
8 Permeability of Lufkin Noncalcareous Smectite with .01N $CaSO_4$ 26
9 Permeability of Calcareous Smectite with .01N $CaSO_4$. . . . 27
10 Permeability of Kaolinite Clay with .01N $CaSO_4$ 28
11 Permeability of Hoytville Illite with .01N $CaSO_4$. 29
12 Permeabilities by site and Unified Soil Classification. . . 32
13 Permeability of Kaolinite Clay with methanol and .01N $CaSO_4$. 36
14 Permeability versus molding water content 37
15 Permeability of Calcareous Smectite vs. molding water content 39
16 Confounding of treatment effects with molding water content . 40
17 Confounded molding water content and gradient effects . . . 41
18 Permeabilities of Faceville Clay after 321.75 hours. . . . 43
19 Permeability of Faceville Clay with .01N $CaSO_4$ (Test #25). . 45
20 Initital permeability of Faceville Clay. 46
21 Permeability of Faceville Clay. 47
22 Water control tests for Faceville Clay 48
23 Increase in Log$_{10}$(K) of Faceville Clay between 321.75 and 1070.9 hours 49
24 Permeability of Faceville Clay (Comparison of rigid-wall and flexible-wall permeameters) 51
25 Long-term test results on Faceville Clay with .01N $CaSO_4$ . . 53
26 Water control test of Hoytville Illite with .01N $CaSO_4$. . . 55
27 Permeability of Elgin Kaolinite with .01N $CaSO_4$ (flexible) . 56
28 Permeability of Elgin Kaolinite with .01N $CaSO_4$ (rigid) . . 57
29 Permeability of Kaolinite Clay with distilled water (rigid) . 58
30 Permeability of Hoytville Illite with .01N $CaSO_4$. 60
31 Decrease in permeability of a flexible-wall sample with water 63
32 Permeability of Kaolinite Clay with methanol (flexible) . . 64
33 Apparent effect of consolidation on a flexible-wall sample with water. 65
34 Flexible-wall tests on Kaolinite Clay with .01N $CaSO_4$ and methanol 67
35 Permeability of Kaolinite Clay with methanol (flexible) . . 70
36 Permeability of Kaolinite Clay with .01N $CaSO_4$ (flexible). . 71
37 Flexible-wall tests on Kaolinite Clay with .01N $CaSO_4$ and methanol with gradients of 2, 10, 30, 50, 100, and 300 . . 72

Number Page

38 Permeability of Noncalcareous Smectite with .01N $CaSO_4$ (flexible) 74
39 Permeability of Lufkin Noncalcareous Smectite with methanol (flexible) 75
40 Effect of gradient on permeability of Lufkin Noncalcareous Smectite with methanol (flexible) 76
41 Hoytville Illite with methanol (flexible) 77
42 Hoytville Illite with .01N $CaSO_4$ (flexible). 78
43 Effect of gradient on the permeability of two clays 79
44 Effect of gradient on the permeabilities of Kaolinite Clay and Micaceous Clay with .01N $CaSO_4$ (rigid-wall tests). . . 81
45 Kaolinite clay with acetic acid in rigid-wall and flexible-wall permeameters. 83
46 Kaolinite clay with trichloroethylene in rigid-wall and flexible-wall permeameters. 84
47 Permeability of Kaolinite Clay with heptane in rigid-wall and flexible-wall permeameters 86
48 Permeability versus number of pore volumes of flow for Kaolinite Clay permeated with methanol. 87
49 Permeability of Lufkin Noncalcareous Smectite with heptane. . 88
50 Permeability of Hoytville Illite with heptane 89
51 Permeability of Lufkin Noncalcareous Smectite with methanol . 90
52 Permeability of Hoytville Illite with methanol. 91
53 Permeabilities of Kaolinite Clay in flexible- and rigid-wall permeameters at various concentrations of methanol. . . . 93
54 Permeability of Kaolinite Clay with methanol (rigid). . . . 95
55 Kaolinite Clay with methanol (flexible) 96
56 Permeabilities of Kaolinite Clay in rigid- and flexible-wall permeameters at various concentrations of acetic acid. . . 97
57 Kaolinite Clay with acetic acid (rigid-wall permeameters) . . 98
58 Kaolinite Clay with acetic acid (flexible-wall permeameters) . 99
59 Permeability of Faceville Clay (Comparison of rigid- and flexible wall tests). 100
60 Permeability of Whitestore Montmorillonite (Comparison of rigid- and flexible-wall tests) 101
61 Permeability of Hoytville Clay (Comparison of rigid- and flexible-wall tests). 102
62 Effect of molding water content on the permeability of Hoytville Illite (with .01N $CaSO_4$) 104
63A-63B Permeability of nonsaturated Elgin Kaolinite with acetone . .108
64A-64B Permeability of presaturated Elgin Kaolinite with acetone . .109
65A-65B Permeability of nonsaturated Elgin Kaolinite with xylene . .110
66A-66B Permeability of presaturated Elgin Kaolinite with xylene . .111
67A-67B Permeability of nonsaturated Micaceous Clay with acetone . .112
68A-68B Permeability of presaturated Micaceous Clay with acetone . .113
69A-69B Permeability of nonsaturated Micaceous Clay with xylene. . .114
70A-70B Permeability of presaturated Micaceous Clay with xylene. . .115
71A-71B Permeability of nonsaturated Bentonite Clay with acetone . .116
72A-72B Permeability of presaturated Bentonite Clay with acetone . .117
73A-73B Permeability of nonsaturated Bentonite Clay with xylene. . .118

Number		Page
74A-74B	Permeability of presaturated Bentonite Clay with xylene. .	.119
75	Permeability of presaturated Elgin Kaolinite with xylene (Tests 38 and 39).	.121
76	Clay soil with exponentially increasing permeability. . .	.123
77	Permeability of presaturated and nonsaturated Bentonite Clay with xylene.	.124
C.1	Simulated Test Data Set 1 for the equilibration problem. .	.170
C.2	Posterior probability distribution of the time at which equilibrium is reached (Data Set 1).	.175
C.3	Simulated Test Data Set 2 for the equilibration problem. .	.176

TABLES

Number		Page
1	Descriptive statistics of permeabilities.	16
2	Descriptive statistics of $\log_{10}$ permeabilities.	17
3	Soil groups	31
4	Incomplete block design--an example	135
A.1	Tabulation of $\log_{10}(k)$ versus k.	143
B.1	Analysis of variance--Completely randomized design . . .	146
B.2	Analysis of variance--Randomized complete block design . .	147
B.3	Analysis of variance--Two-way treatment structure--Completely randomized design.	148
B.4	Analysis of variance--Two-way treatment structure replicated in randomized complete blocks.	149
B.5	Analysis of variance table--Incomplete block design . . .	158
B.6	Analysis of variance table--Split-plot design	160
B.7	Split-plot design with two whole plot factors and one sub-plot factor	162

ACKNOWLEDGMENTS

The writers wish to acknowledge the assistance, helpful suggestions, and contributions of Jonathan Herrmann, the technical project monitor, and Walter E. Grube, Jr., Daniel Greathouse, and Lewis Rossman, all of the EPA.

The contributions and assistance of Kirk W. Brown, David E. Daniel, the Ohio River Division Laboratory of the U.S. Army Corps of Engineers, and Jeffrey J. Peirce in providing the data for this project are gratefully acknowledged.

The assistance of Kate Allen, Jim Albertz, and Hsu Chyi-Hung in obtaining and processing the data is also appreciated.

SECTION 1

INTRODUCTION

Clay liners and clay caps are important components of many existing hazardous waste land disposal facilities. Clay liners are also specified in the Hazardous and Solid Waste Amendments of 1984 and must exhibit a permeability of no more than 1×10^{-7} cm/sec (42). In engineering practice the permeability of proposed clay liner material is customarily determined in geotechnical engineering laboratories using water as the permeant. A few years ago, however, it became evident from the research of Brown and Anderson (6) that certain chemicals have pronounced effects on clay liner permeabilities. These permeabilities were determined using laboratory permeability apparatus.

The Land Pollution Control Division (LPCD) of the Hazardous Waste Engineering Research Laboratory for the United States Environmental Protection Agency is in charge of investigating the effectiveness of clay liners in preventing the seepage of chemical leachates and contaminated cut-off water from hazardous waste containment facilities. Since the findings of Brown and Anderson, LPCD has sponsored additional research on the permeabilities of various clays. Of particular interest was the immediate and long-term permeability of clay liner materials when subjected to hazardous contaminants and leachates.

These investigations revealed that certain chemicals may indeed have dramatic effects on compacted clays. However, considerable variability was observed in the equilibrium permeabilities from the various laboratories and individual researchers. There may be several reasons for this. Some investigators used rigid-wall permeameters, while others used flexible-wall permeameters. Further sources of variability may have been the varying percentages of molding water contents, the individual testing techniques used, the methods of deairing the samples and the degree of water-saturation achieved before the introduction of chemicals, the type and concentration of chemicals used, and the various driving gradients applied to the test specimens.

Having access to a considerable quantity of accumulated data on clay liner permeabilities, LPCD was interested in establishing the overall reliability of the laboratory measurements of clay liner permeabilities and the reasons for permeability measurement variability. For this reason, the present study was undertaken with the following objectives:

a) To collect available historic data on permeability measurements from both public and private sources and to identify the test methods and parameters used.

b) To analyze the collected data in order to determine the degree of variability likely to be found among the results of permeability tests and to identify sources of errors in the testing procedures and estimate their significance.

c) To make recommendations to enhance the applicability of statistical methods to the analysis of future clay liner permeability tests.

The co-principal investigators conducted the work with assistance from graduate and co-op students. The primary work station was the University of Cincinnati, with some tasks performed at the Center Hill Research Facility. All work was coordinated with Mr. Jonathan Herrmann, EPA Technical Project Monitor (TPM). This work was performed during the period from April 1, 1984, to September 30, 1985, in accordance with the terms of Contract No. 68-03-3210-03, under the provisions of the "Additional Work-As Needed" section thereof.

SECTION 2

CONCLUSIONS AND RECOMMENDATIONS

The following points will be discussed, and examples given, in the remainder of the report.

a) Given that they have achieved equilibrium, the variability of permeability test results using standard aquaeous leachate, standardized compaction, and samples prepared from a uniformly moisturized batch of homogeneous clay soil appears to be reasonably small relative to other sources of error. The variability of tests using chemicals as permeants differed substantially, with some clay-chemical combinations yielding variability characteristics comparable to water permeation and others resulting in extreme inconsistency.

b) Certain series of tests in the collected data indicated that, even with water as the permeant liquid, reaching liquid flow equilibrium may be somewhat more difficult to establish than is commonly believed. Furthermore, the error in a permeability determination which may be caused by premature termination may not be small in comparison to the variability of the initial permeabilities of replicated tests. This implies that the length of testing time commonly found in private sector tests in the collected data may be insufficient.

   Because the measured permeability of a test often undergoes a prolonged period of transitory behavior before leveling off at an equilibrium value, deciding when a test has exhibited stable permeability long enough to conclude that equilibrium has been achieved is not straightforward. Development of objective criteria, which may be useful in confirming the experimenter's judgement in this decision, is a topic which deserves considerable attention in future permeability research. In Appendix C, two approaches are described which are currently under investigation.

   The use of identically prepared and concurrently run water control tests is recommended in comparative tests of chemical permeants, to reduce ambiguity in distinguishing between the effects of equilibration and chemical.

   In some situations in which chemical permeants have strong effects on the clay samples and lead to channeling and crack formation which in turn lead to massive increases in conductivity, it may be that the concept of specimen equilibration is less useful than measures based on the notion of time to failure. If this is the case, then there is a need for the development of practically meaningful failure criteria for use in research studies.

c) The permeability of a soil sample is extremely sensitive to the parameters of its preparation, such as the molding water content at which it is compacted. This appears to be true not only for water as permeant but for chemical permeants as well. Minor variations in these parameters can cause changes in permeability which may be large relative to the accuracy of the test itself.

Methods of statistical experimental design should be applied in comparative permeability research experiments to control the effects of factors such as molding water content. These methods, which include randomization and blocking, are discussed in Section 5.2.

The sensitivity of permeability to factors such as molding water content also indicates a need for research which would estimate the bounds within which these factors can be maintained during the actual construction of a liner. If these factors vary substantially during liner construction, the results of tests performed only at their target values may give little indication of the future performance of the liner.

d) Several series of tests on which data have been provided appear to indicate that the gradient at which tests are run may have an effect on the resulting permeabilities in both rigid-wall and flexible-wall tests, the magnitude of which in some cases appears to be practically significant. Such a phenomenon could be caused by factors whose effects are opposite in nature. That is, increased gradient could cause sample consolidation leading to decreased permeability, which appears to have occurred in several series of tests discussed in this report. In tests of chemical permeants in rigid-wall permeameters, on the other hand, there are data that suggest that higher gradients may promote channel formation and prevent clogging of flow passages.

e) Other variations in testing methods may be expected to affect test results as well. In the tests of a bentonite clay using xylene as the permeant in rigid-wall peameameters, Brown obtained initial permeabilities in nonsaturated tests that were four orders of magnitude greater than those found in corresponding tests which were presaturated by permeation of .01N $CaSO_4$ prior to addition of the xylene. Even after two or more pore volumes of flow, the permeabilities of the presaturated tests had not reached the level of the nonsaturated tests.

f) There were almost no within-lab comparisons of flexible-wall and rigid-wall tests of concentrated chemical permeants in this data. The data that are available, together with between-lab comparisons under admittedly dissimilar conditions, suggest that often quantitatively different results may be obtained by the two methods. They also suggest that flexible-wall test behavior may be somewhat less erratic in testing certain chemicals.

g) It is important to assess the degree of uncertainty of permeability determinations due to the nature of the test itself. It is equally important to estimate the variability of permeabilities that might occur in location-to-location sampling from within a proposed source of liner material in order to judge the amount of sampling required to ascertain the acceptability of the soil within the source. Limited data from the private sector suggest that permeabilities of samples taken from various positions within a site exhibit far more variability than do the replicate tests on uniform clays observed in the research setting, even when the samples are relatively homogeneous with respect to Unified Soil Classification, Atterberg limits and grain size information. Although it seems reasonable to ascribe this variability in large part to location-to-location heterogeneity of the soil, for a number of reasons discussed in the text of the report this factor cannot be completely separated in the data from variability which has been caused by differences in sample preparation and testing. Additional research in this area is indicated.

SECTION 3

METHODS AND PROCEDURES FOR DATA COLLECTION AND ANALYSIS

3.1 COLLECTION AND ORGANIZATION OF DATA

Selection of Data Sources

Permeability data were to be collected from laboratories that were active either in research or design of clay liners. Nine data sources were identified based on established contacts, past and on-going research, data availability and publications. These included four EPA contract sources, two government laboratories, two private consulting firms, and one independent university source. The EPA contract sources were Brown and Anderson of Texas A&M University, College Station, Texas, D. Daniel of the University of Texas, Austin, Texas, J. Peirce of Duke University, Durham, North Carolina, and Y. Acar of Louisiana State University, Baton Rouge, Louisiana. The two government laboratory sources were the Ohio River Division Laboratories of the Corps of Engineers, U.S. Army, Mariemont, Ohio, and the Illinois State Geological Survey, Champaign, Illinois. Two private consulting firms and their geotechnical laboratories were selected and are identified as companies ABC and XYZ. They chose to remain anonymous to protect their clients. The independent university laboratory source was R. Stephenson of the University of Missouri-Rolla, Rolla, Missouri.

Collection of Limited Data Set

The above potential data sources were contacted to determine what kind of permeability data they had, the approximate number of tests they had conducted on clay liner materials, and whether their permeability data were available for this project. A survey sheet was filled out for each source to summarize his response. All nine sources contacted expressed their willingness to cooperate. After that, a written request was made to each of the nine sources to provide all available and suitable raw permeability data, including clay characterization and test specification data, and any available reports on these tests.

Data sets were received from Brown and Anderson, Daniel, Peirce, the Ohio River Division Laboratories of the U.S. Army Corps of Engineers Laboratory, Company ABC and Company XYZ. Published reports and theses were received from the other three sources, but no raw data.

Description of Data

Two separate sets of data were received from Brown and Anderson (6), (7). The first set contained data on 32 tests, and it tabulated the permeability test results as pore volume flows versus the coefficients of permeability. A report on these tests was also received. The second set presented data on 67 tests. Again, for each test the pore volume flows versus the coefficients of permeability were tabulated. No report accompanied this set of data. However, information on the clays used and the test parameters was supplied by the researchers. Brown and Anderson's data contained tests on Calcareous Smectite, Hoytville Illite, two Kaolinites, Lufkin Noncalcareous Smectite, a Micaceous Clay, and a Bentonite. The permeants used were: water, acetic acid, aniline, ethylene glycol, acetone, methanol, xylene and heptane. Various reports and articles by Brown and Anderson were also available for reference.

Three sets of data were received from Daniel (14). The first set consisted of data on 103 permeability tests. These were run on a Kaolinite, Lufkin Noncalcareous Smectite and Hoytville Illite, using primarily water and methanol as permeants. A few tests using heptane were also included. The second set had permeability data on 30 tests. These were all conducted on a Kaolinite with various concentrations of methanol, acetic acid, heptane and trichlorethylene. Permeability data on 6 tests run on Dow Illite made up the third set. These were tested with various concentrations of methanol. The data sheets each had a tabulation of pore volumes of flow versus permeability. Also included was information on the permeameter type, soil type, compaction water content, dry unit weight, the type of permeant liquid and the hydraulic gradient used. In addition, a report and several articles were provided by Daniel.

Permeability data on 38 tests were received from Peirce (40). These tests were conducted on three clays: Whitestore Montmorillonite, Hoytville Illite and Faceville Kaolinite. The permeants used were: water, ferric chloride and nickel nitrate. The test data sheets tabulated cumulative time and cumulative pore volume flow versus permeability. The permeability data were accompanied by basic classification data on the clays tested. Three quarterly reports were also received from Peirce.

From the Ohio River Division Laboratories, permeability data were received for 66 tests (38) on various soils. The data were transmitted on copies of laboratory test sheets. These contained the permeability measurements and plots of void ratio versus permeability.

Data on 45 separate permeability tests were received from Company ABC (11) on various soils. These consisted of copies of the laboratory data sheets. They contained time and flow measurements for each test and the calculated coefficients of permeability. The data sheets also included information on sample sizes, moisture contents, and wet and dry densities of the samples tested. Separate data sheets provided information on other pertinent soil properties, such as optimum moisture content and maximum density, when available.

Data on 30 laboratory permeability tests on various soils were obtained from Company XYZ (12). These consisted of tabulated data on elapsed test time, flow volume and calculated permeability. In addition, data on sample sizes and classification of the soils were also furnished.

Organization of Data

The permeability data received were organized by filling out a "Permeability Test Specification Sheet" (Figure 1) for each test, and attaching it to its corresponding permeability test data sheets. In addition, for each clay for which permeability data were received, a "Clay Type Sheet" (Figure 2) was filled out and was also attached to the test data sheets.

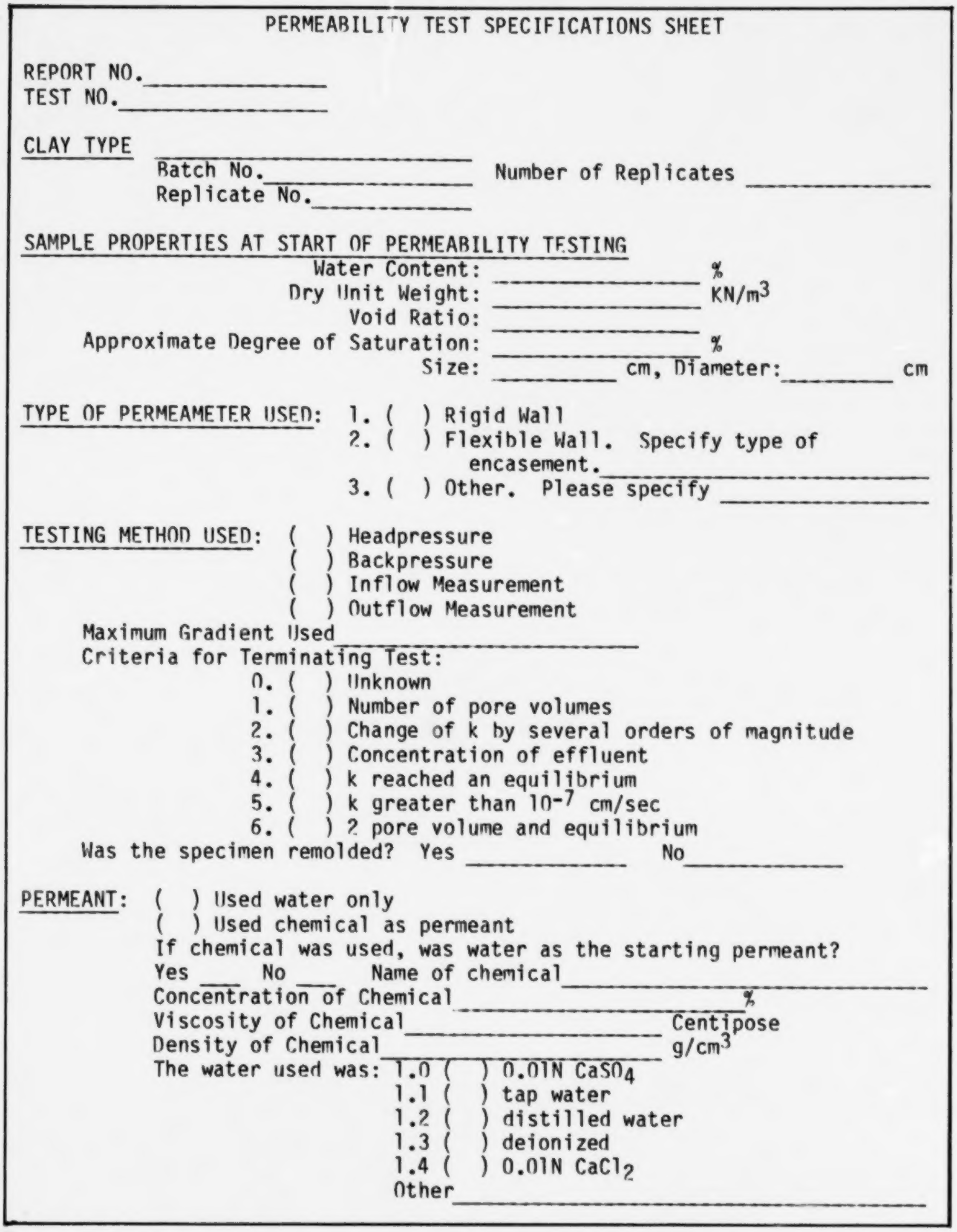
The Clay Type Sheet contains a letter designation for the clay type used, information on the site from which the clay was obtained, a description of the clay, classification data such as liquid limit, plasticity index, Unified Soil Classification, textural classification, mineralogical description, organic content, pH, color, cation exchange capacity, specific gravity, shrink-swell capacity, Standard Proctor Density and optimum moisture content (see Figure 2).

The typical Permeability Test Specification Sheet (Figure 1) contains the following items:

- Clay type designation as it appears on the corresponding Clay Type Sheet and a test designation number.
- Sample properties at the start of the permeability test, such as water content, dry unit weight, void ratio, degree of saturation, size of sample, and whether the sample is undisturbed or remolded.
- Type of permeameter used, such as rigid-wall, flexible-wall, or consolidometer type, and type of encasement if flexible-wall.
- Testing method used, such as pressurized head only, pressurized head and backpressure, etc., measurement of both inflow and outflow, maximum gradient used, criterion for terminating the test, and the number of replicates tested.
- Information on permeants used, such as: was water used only; was a chemical used, and if so, was water used as the starting fluid; the name of the chemical, its properties, and concentration; and the characteristics of the water used.

3.2 TEST METHODS AND EQUIPMENT USED BY DATA SOURCES

The laboratories that supplied data and information used various test techniques. The constant-head type permeability test arrangement was used by the following researchers and laboratories: Brown and Anderson, Company XYZ, and Daniel. The falling head permeability test was used by Company ABC and



PERMEABILITY TEST SPECIFICATIONS SHEET

REPORT NO.____________
TEST NO.______________

CLAY TYPE ____________________
Batch No._____________ Number of Replicates ____________
Replicate No.___________

SAMPLE PROPERTIES AT START OF PERMEABILITY TESTING
Water Content: ______________ %
Dry Unit Weight: ______________ KN/m^3
Void Ratio: ______________
Approximate Degree of Saturation: ______________ %
Size: ________ cm, Diameter:________ cm

TYPE OF PERMEAMETER USED:
1. ( ) Rigid Wall
2. ( ) Flexible Wall. Specify type of encasement.____________________
3. ( ) Other. Please specify ______________

TESTING METHOD USED:
( ) Headpressure
( ) Backpressure
( ) Inflow Measurement
( ) Outflow Measurement

Maximum Gradient Used____________________
Criteria for Terminating Test:
0. ( ) Unknown
1. ( ) Number of pore volumes
2. ( ) Change of k by several orders of magnitude
3. ( ) Concentration of effluent
4. ( ) k reached an equilibrium
5. ( ) k greater than 10^{-7} cm/sec
6. ( ) 2 pore volume and equilibrium

Was the specimen remolded? Yes __________ No__________

PERMEANT:
( ) Used water only
( ) Used chemical as permeant
If chemical was used, was water as the starting permeant?
Yes ___ No ___ Name of chemical____________________
Concentration of Chemical __________________%
Viscosity of Chemical_________________ Centipose
Density of Chemical___________________ g/cm^3
The water used was:
1.0 ( ) 0.01N $CaSO_4$
1.1 ( ) tap water
1.2 ( ) distilled water
1.3 ( ) deionized
1.4 ( ) 0.01N $CaCl_2$
Other____________________________

Figure 1. Permeability test specifications sheet

CLAY TYPE SHEET

(Please fill out a separate sheet for each clay tested)

Clay Type : ________________

(This is a relative designation for this questionnaire. For each type of clay you have worked with, please assign a letter. For example, if you have worked with 5 different clays, you will report clay type A on one sheet, clay type B on a second sheet etc.)

Site (Excavation Location, Source etc.):

Description:

Classification: LL= ____
PI= ____
Grain Size Curve (Please Attach)
% of Clay ______, % of Silt ______, % of Sand ______

Unified Classification: ____

Textural Classification: ____

Mineralogical Description: ______________________________________
____

Organic Carbon Content: ____________ %

pH: ____________

Color: ____

Optimum Moisture Content ______ % Max. Dry Unit Weight _____ kN/m^3 _____
(Standard Proctor)

Cation Exchange Capacity ________ meg/100g

Specific Gravity _________ Shrink-Swell Capacity ____________

Figure 2. Clay type sheet

Peirce. Consolidometer type permeability tests were run by the Ohio River Division Laboratories.

Three types of permeameters were used by the data sources. Brown and Anderson, Daniel, Peirce and Company XYZ used the so-called rigid-wall type permeameter. For their tests they used the standard compaction mold for the preparation and encasement of the permeability specimens. Company ABC, and also Daniel and Pierce as an alternate approach, used flexible-wall permeameters. To use the flexible-wall permeameter, the soil samples first had to be compacted in a rigid mold, then removed, trimmed and encased in a flexible membrane and placed in a triaxial chamber for lateral confinement. The ordinary membrane used was latex rubber. However, for chemical leachates a chemically resistant teflon^R tape was usually added between the membrane and the sample.

The consolidometer-type permeameter was used by the Ohio River Division Laboratories. This is a less widely used type permeameter which uses one of the standard consolidation test apparatuses and can also be considered a rigid-wall type permeameter.

3.3 DATA BASE LIMITATIONS

It was known at the initiation of the project that the data base resulting from the data collection effort would be limited in certain respects. The amount, degree of documentation, and structure of data to be provided by the available sources was in most cases largely unknown, and in general, test programs were not anticipated to have been designed with a specific view towards a statistical assessment of the variability of test results. The following seven points summarize important limitations of the available data:

a) Data sources were identified based on the knowledge that they possessed relevant and available data on permeability testing. The selected sources were in no sense chosen randomly from any pool of available sources, and there is no certainty that the data so obtained is typical of data that might have been obtained elsewhere

b) For many of the tests obtained from private sources, concomitant information concerning the preparation of specimens and the physical characteristics of the clay soils was not as complete as would be desired. Precise data on the location from which samples were taken was in some cases not available. There was virtually no replication of tests from within the same samples. All private source tests used only water as a permeant

c) As expected, data received from EPA contract sources was highly project specific in the sense that researchers in general used different types of equipment and test methods, and investigated different clay-chemical combinations. Data obtained from Brown and Anderson used rigid-wall permeameters exclusively; tests made available by Peirce used rigid-wall permeameters with the exception of six tests, one on each combination of three clay soils and two permeants; results from Daniel used the flexible-wall apparatus primarily, although a number of rigid-wall tests were provided

d) In some cases, no replications of tests on specific clay-permeant combinations were conducted

e) Criteria for test termination varied considerably, sometimes making the assignment of a terminal permeability rather subjective. Often, tests were terminated prior to the point in time where equilibrium was clearly established. In situations where chemical permeants lead to sidewall erosion, channel formation, cracking, or silting and subsequent clogging, the notion of equilibrium becomes questionable. This leads to extreme complications in attempting to reasonably quantify the degree of variation between tests

f) As a general rule, precise information on the level at which tests were replicated was not available. Central to this point is the question of whether replicate specimens were prepared from a single moisturized batch of clay soil, or from separately moisturized batches. In the former case, less variability of test results would be expected, due to presumably tighter control of compaction water content as well as improved homogeneity of the clay soil and specimen preparation. In no case were tests systematically replicated at both the within and between-batch levels, as would be required for a rigorous analysis of the components of variance in test results to be attributed to the operations of moisterization, compaction and testing

g) In general, there appeared to be little conscious effort to randomize experimentation with regard to either the assignment of experimental material to treatments or time sequence of testing.

The limitations described above impose constraints on the types of analyses which are feasible. Between-lab comparisons are difficult and easily subject to incorrect attempts at generalization due to the confounding effect of extraneous factors. Data for within-lab comparisons are often lacking or too limited to form the basis for generalization. Generally speaking, the most useful analyses consisted of simple exploratory methods of a graphical nature. For this reason, the discussion to follow should be regarded not as a description of the results of formal tests of pre-specified hypotheses or rigorous methods of estimation, but rather as a summarization of the data at hand and the general patterns that appear in that data.

Despite the above limitations, it is believed that much has been gained from the analysis. In some cases, trends have been observed that confirm previous work reported in the literature, but which may well be unfamilar to many persons who must make decisions on the basis of permeability data. In other cases, trends were discovered which sometimes were surprising to these writers, and which may be used as a basis for hypothesis formulation, useful in the design of future experimentation. Factors which apparently influence the variability of test results in the data have been identified, and very approximate indications of the degree of this variability have been noted. Finally, general remarks can be made regarding statistical considerations relevant to promoting the usefulness of methods of statistical analysis for permeability test data.

3.4 COMPUTER HARDWARE AND STATISTICAL ANALYSIS PACKAGES

All computations performed in the analyses described in this report were performed at the University of Cincinnati, Cincinnati, Ohio on either an Amdahl 470 V/7A or an IBM 3081D computer. Data transformations and summarizations were accomplished by use of SAS (27) arithmetic capabilities, built-in functions or data summarization procedures. All statistical analyses performed or discussed in this report are available through SAS procedures or in the BMDP statistical software computer programs (18), with the exception of those discussed in Appendix C. These analyses require the construction of programs which may be written either in a procedural language such as FORTRAN or by use of the programming capabilities provided within SAS PROC MATRIX.

SECTION 4

DATA ANALYSIS RESULTS AND DISCUSSION

4.1 SCALES OF MEASUREMENT AND VARIABLILITY

Throughout this report the magnitude of the variability observed among the results of laboratory permeability tests and the factors which appear to contribute to that variability are discussed. In order to report results which have any consistency across the series of tests which were provided for analysis and which may be generalized to some degree, it is necessary to provide numerical summaries, such as standard deviations, which are based on a measurement scale in which the degree of variability is as homogeneous as possible for different clay soils and different test conditions. Considering the extreme range over which the permeability of soils may vary, this selection of a scale of measurement is critical. In our data set are samples whose permeability to water alone ranges over six orders of magnitude.

The measurement scale used in this report is logarithmic (the base 10 is used throughout). Selection of this metric was made primarily on an empirical basis. That is, permeability measurements on this scale exhibited far greater homoscedasticity (that is, constant variance) across clay types than did the raw permeabilities themselves, and no other transformation was identified that was consistently superior to the logarithmic in this regard.

Additionally, there are several reasons why the logarithmic scale is in some sense the "natural" choice of metric for describing the variability of permeability measurements, as will be discussed below. It was also found that use of this transformation simplified the relationship between permeability and certain related quantities, such as void ratio in consolidometer type permeability tests.

Finally, a logarithmic scale is meaningful to those who work with permeability data, since graphical displays of such data are virtually always presented on semi-logarithmic paper, which visually makes the transformation from raw permeability to log permeability.

To motivate use of the logarithmic scale, consider Figure 3, which shows the permeability of four different clay soils to water (.01N $CaSO_4$). These data are from Brown and Anderson (6); the soils in question are Calcareous Smectite, Hoytville Illite, a Kaolinite and Lufkin Noncalcareous Smectite. Altogether there are 43 permeabilities illustrated in Figure 3. These tests were compacted into rigid-wall permeameters with nominally equal compactive efforts and within fairly narrow limits on molding water content for each clay type. The control on molding water content was sufficient to prevent a statistically significant relationship between permeability and measured molding water content for any of the soils (see Section 4.3).

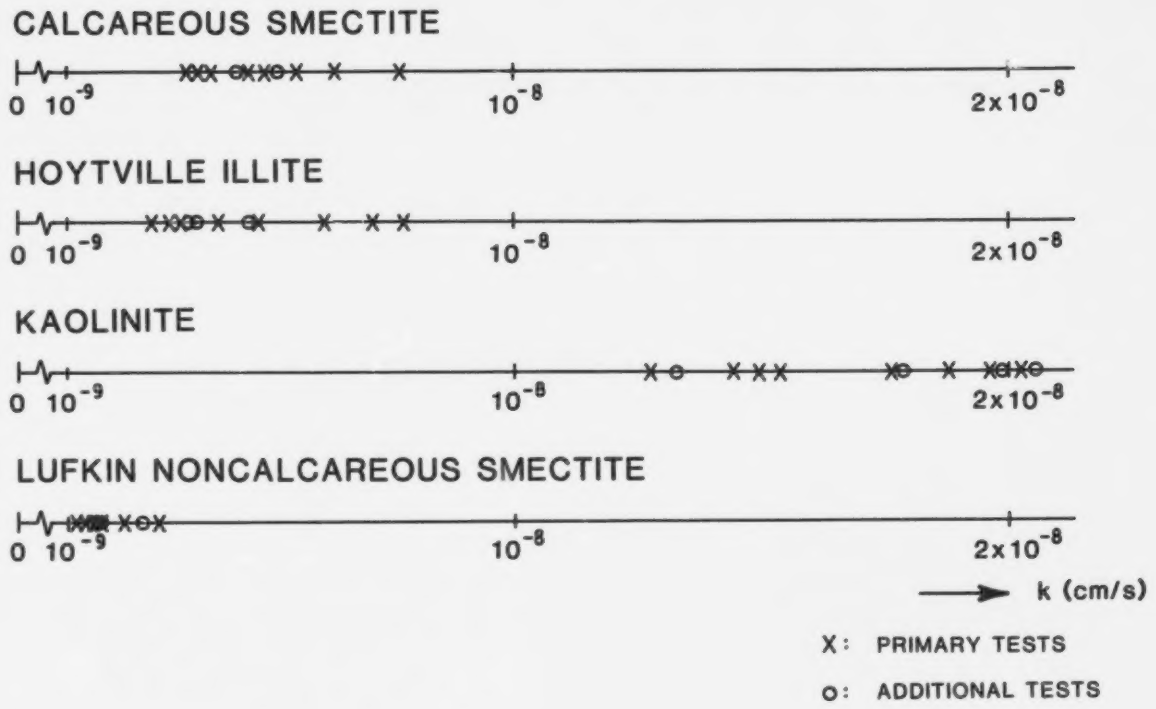


Figure 3. Permeabilities of four clays with .01N $CaSO_4$ water

In Figure 3 "X's" represent tests that were included in the data set and are based on initial water permeation of 32 specimens which were later subjected to various chemical leachates in Brown and Anderson's comparative study. The remaining eleven tests, the results of which are denoted by "O's", were obtained from Tables A-1 to A-4 of Brown and Anderson's report (6), which summarize the results of 25 additional water tests used to estimate the sensitivity of permeabilities of these clays to molding water content. The additional tests selected for use here are those compacted at nominally the same molding water content as the remaining specimens.

The duration of the tests is short in terms of total pore volumes of flow, which varies from .1 to .56 pore volume (in terms of time, however, the longest of the tests on the relatively impermeable Lufkin Noncalcareous Smectite required about two months). Therefore, the permeabilities shown in Figure 3 are based on the total flow for each test.

The most noticeable feature of Figure 3 is that the variation of permeabilities differs considerably across the four clay soils, with the permeabilities of the Kaolinitic specimens about five times more spread out than those of the samples of Lufkin Noncalcareous Smectite. Further, the permeabilities in most cases appear to be somewhat skewed to the right. The pronounced difference in variability is also evident from the summary data in Table 1, as well as from the results of statistical tests for constant variance. Both Bartlett's and Levene's tests (30), (34) are significant at the .001 level.

TABLE 1 DESCRIPTIVE STATISTICS OF PERMEABILITIES

Clay	Number of Specimens	Mean Permeability (cm/sec)	Standard Deviation (cm/sec)	Coefficient of Variation	Test For Normality (Shapiro-Wilk) p-value
Calcareous Smectitle	10	5.00×10^{-9}	1.37×10^{-9}	27.3	.43
Hoytville Illite	11	4.62×10^{-9}	1.71×10^{-9}	37.1	.21
Kaolinite	12	1.72×10^{-8}	2.81×10^{-9}	16.4	.22
Lufkin Noncalcareous Smectite	10	1.77×10^{-9}	5.64×10^{-10}	31.8	.04

There is a clear relationship between the standard deviation of permeabilities within soil types and their mean permeability, with the degree of variability increasing as the mean increases. Basically, the relative variability of permeabilities is much more stable across clay types than is the absolute level of variability. This phenomenon has been observed in all of the data which have been investigated. If relative variability is roughly constant across clay types, then a logarithmic transformation of the permeabilities should result in a greater degree of homogeneity of variance.

In Figure 4, the data are shown after re-expression by a base 10 logarithmic transformation. Table 2 gives corresponding numerical summaries.

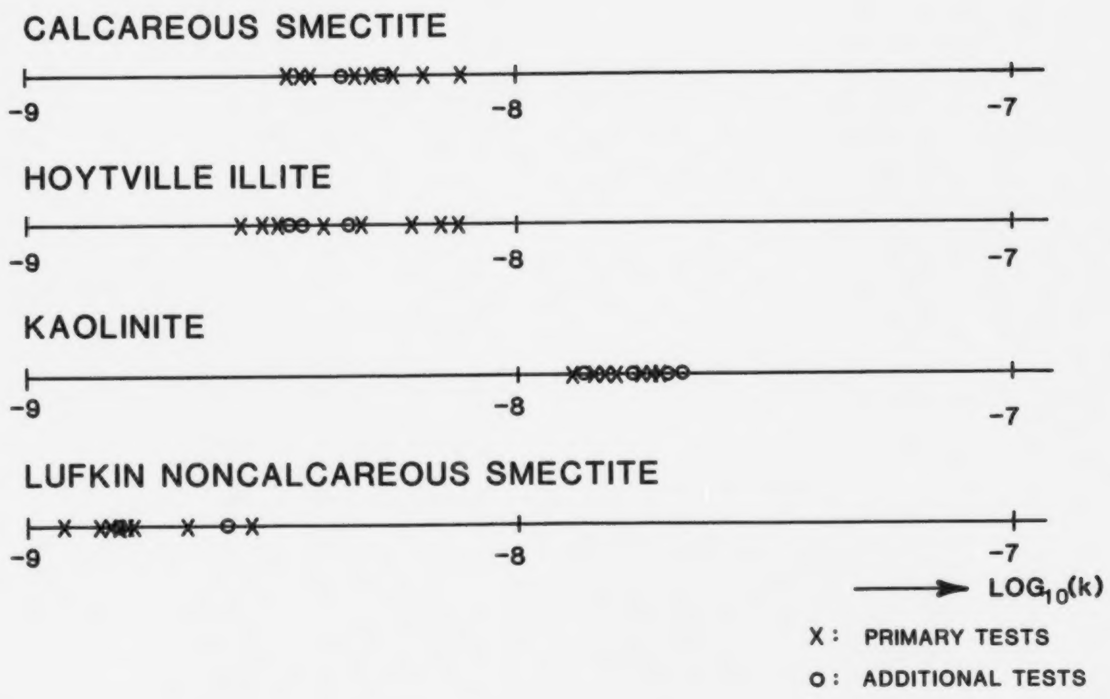
TABLE 2 DESCRIPTIVE STATISTICS OF LOG_{10} PERMEABILITIES

Clay	Number of Specimens	Average Log Permeability	Standard Deviation	Test For Normality (Shapiro-Wilk) p-value
Calcareous Smectitle	10	-8.31	0.11	.66
Hoytville Illite	11	-8.36	0.15	.46
Kaolinite	12	-7.77	0.074	.18
Lufkin Noncalcareous Smectite	10	-8.77	0.13	.18

The degree of variation is clearly more uniform across clay types on a logarithmic scale of measurement, as is evident both from the standard deviations in Table 2 and the graphs of Figure 4. On this scale, however, the Kaolinite measurements appear to be somewhat less variable than the others. While formal statistical tests of constant variance do not indicate that it is entirely implausible to attribute this to chance (Bartlett's test gives a p-value of .17, whereas Levene's test, which is insensitive to a normality assumption, yields a p-value of .09), there is in fact no good reason to expect that permeability measurements of different clays would exhibit precisely equal variability on any scale of measurement.

There are various exploratory methods for determining a transformation of data to achieve homogeneity of variance (5). When these methods are applied to the data of Figure 3, they point toward a cube-root transformation. However, in other series of tests in the data set, there did not seem to be a consistent indication that this or any other transformation was superior to the logarithmic in stabilizing variance. For this reason, together with the familiarity of the logarithmic scale among users of permeability data, it was decided that it was the most appropriate for this study.

There are also theoretical models which motivate the logarithmic transformation. A number of such models indicate that certain factors which may contribute to the variability of permeability data should have multiplicative effects. For example, changes in temperature affect the viscosity of the permeant liquid, which in theory is inversely proportional (for a given permeant) to the permeability of a clay soil. On the logarithmic scale, the effect is additive, rather than multiplicative. To illustrate, a



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 4. Log_{10} permeabilities of four clays with .01N $CaSO_4$ water

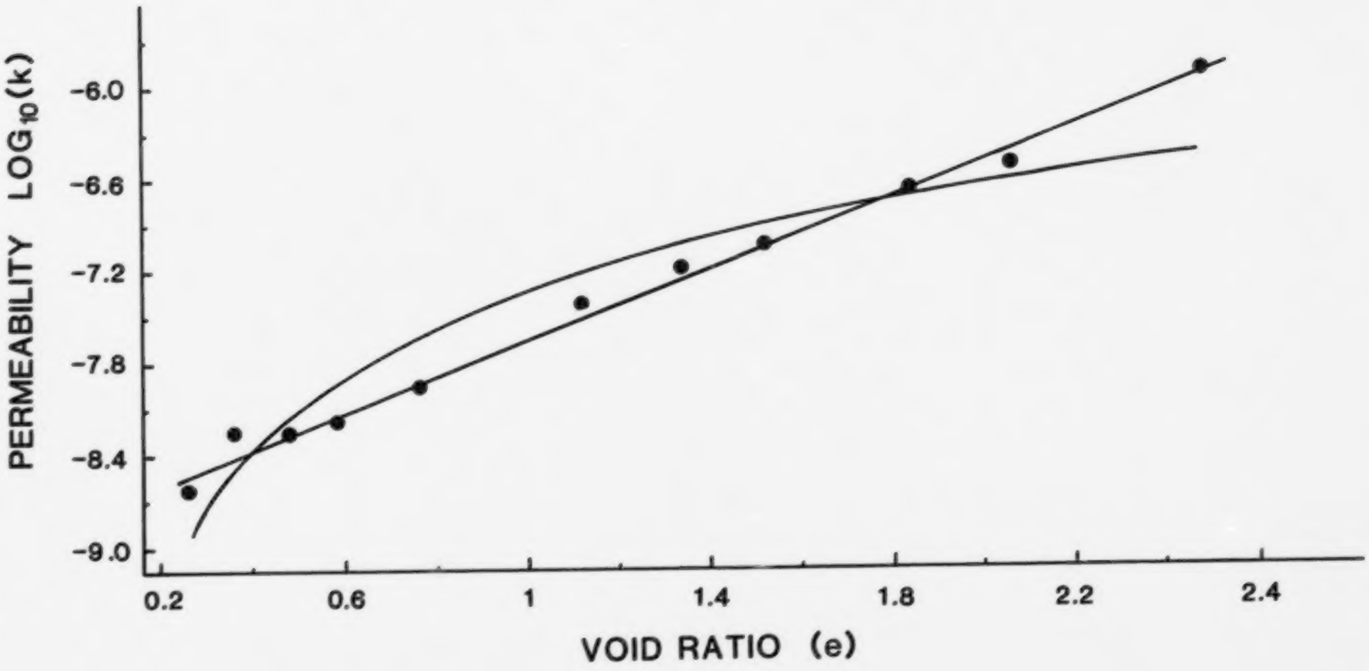
5°C increase in temperature should, at reasonable levels of temperature, decrease the viscosity of water by about 11%. This should have the effect of increasing the permeability of a clay soil by about 12%. If two soils are considered, one having permeability ten times greater than the other, the change in permeability caused by the temperature increase will be ten times as great for the more permeable of the two. On the logarithmic scale, the perturbations caused by the temperature change are additive; the log permeabilities of the two soils should each increase by .05.

A second example has to do with unobserved drifts in the gradient at which a clay is tested. Thus, if the two soils described in the previous paragraph were being tested at the same nominal gradient, a 10% increase in gradient would appear to create a change in measured permeability ten times greater in the more permeable clay than in the less permeable one. On the logarithmic scale, the errors introduced by the gradient drift would both be $\log(1.1) = .04$.

As a final motivation for the choice of a logarithmic scale of measurement, it has been noted that log permeability relates well to certain soil characteristics. For example, data received from the Corps of Engineers displayed the results of consolidometer-type permeability tests by graphing log permeability against void ratio and indicated that this relationship was linear over the ranges of consolidation in question. This point is illustrated in Figure 5, using the results of consolidometer-type permeability tests of Sodium Illite clay over void ratios ranging from 0.25 to 2.3. These data are due to Olsen (39), and were taken from Figure 2.22 of (15), where they were expressed in terms of porosity, rather than in terms of void ratio.

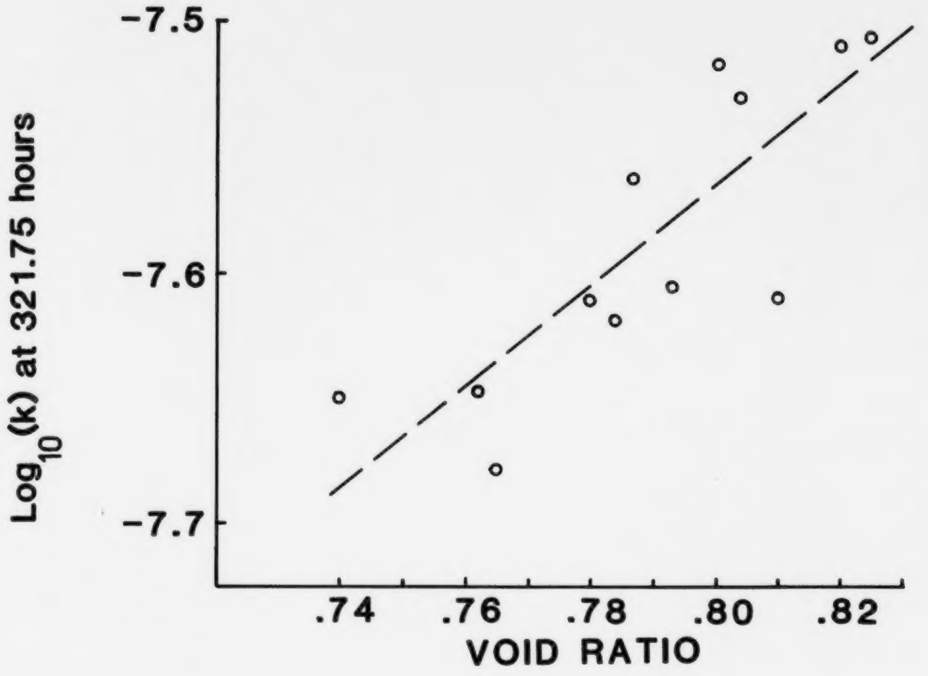
In Figure 5, black dots represent log-transformed permeabilities. The straight line through these points was obtained by regressing log permeability on void ratio. It is clear that the relationship is well represented by a straight line, even over the extreme range of void ratios considered. The linear relationship clearly fits the data better than the curve also shown in the figure. The curve is the result of a least-squares fit of log permeability to the logarithm of $(e^3/1+e)$, where e is void ratio, in accordance with the Kozeny-Carman theory. According to this theory, which works well for sandy soils (15), the regression coefficient of $\log(e^3/1+e)$ should equal one. While the fitted coefficient was almost precisely equal to this (.98), the curve fits the data poorly, indicating a departure from the Kozeny-Carman theory.

In rigid-wall tests, it was also fairly typical to observe a relationship between initial void ratio and log permeability, although this relationship was sometimes obscured by the fact that under standard compaction of replicated samples, the variation in void ratio observed should be small in relation to the accuracy with which it may be measured. Figure 6 shows the results of twelve permeability tests of Faceville Clay due to Peirce, which are discussed in some detail in Section 4.4. Permeabilities of the twelve specimens to .01N $CaSO_4$ are moderately correlated with void ratio ($r = .82$, $p < .01$).



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 5. Effect of consolidation on the permeability of Sodium Illite Clay



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 6. Permeability of Faceville Clay at 321.75 hours (rigid-wall permeameter)

Unfortunately, it is difficult to relate such observed correlation to a single source of variation. While Figure 6 indicates that a reasonable amount of the variability of permeabilities in this series of tests may be due to variation in applied compactive effort, the differences in void ratio could equally well be caused by minor heterogeneity of the soil used in the tests. Furthermore, although the twelve specimens in this series were known to have been moisturized in a common batch, some small sample-to-sample variation in molding water content is inevitable, and the observed differences in void ratio could have been partially caused by the well-known effect of water content on void ratio.

Notice that the scales to which Figures 5 and 6 are drawn are radically different in order to show as much detail as possible. Throughout the report, figures are used extensively, and it is appropriate to warn the reader that the scales frequently vary. Significant scale changes from figure to figure are necessary to provide meaningful illustrations, but can be misleading if not noted.

To conclude this discussion, several general remarks will be made on the degree of variability found in the data made available to these writers, together with some comments on the interpretation of variance estimates on the log scale, which may be of use in reading the remainder of the report.

For reasons outlined in Section 3.3, these writers do not believe that the available data will support a precise analysis leading to accurate variance estimates that may be used to clearly establish the exact precision of various test methods. This would require a coordinated effort leading to experimentation in which the various methods are compared under matched conditions with carefully defined criteria for test termination and permeability measurement, systematic control of such factors as batch effects which might otherwise bias a comparison of methods or variance component estimates, and with careful randomization of the assignment of specimens to methods and the order of specimen preparation. Nevertheless, it is important to summarize the variability observed in the available data in order to provide at least very rough guidelines that may be useful to those who must make decisions based on data provided in permeability tests. Additionally, even an approximate estimate might be valuable to researchers in assessing the degree of variability that might reasonably be expected to occur in future testing, as a guide to planning sufficient replication in their studies.

Almost all of the replicated tests on homogeneous soils in this data set fall into one of the following groups:

a. Tests by Brown, Daniel and Pierce with .01N $CaSO_4$ as the permeant

b. Rigid-wall and flexible-wall tests by Peirce using very dilute $FeCl_3$ and $Ni(NO_3)_2$ solutions as permeant. These permeants had no effect on hydraulic conductivity that was statistically verifiable, and for purposes of this discussion can be considered comparable to water tests

c. Flexible-wall tests by Daniel using methanol as the permeant

d. Flexible-wall tests provided by Daniel using heptane as the permeant; these were very few in number, and were replicated only within one clay, a Kaolinite, for which there were two replicates at a gradient of 100 and five replicates at a gradient of 300

e. Tests provided by Brown in rigid-wall permeameters using concentrated acetone and xylene as the permeant.

The variability of the tests results in groups d and e was very different than those of the other three groups. For clarity of presentation, discussion of group e will be deferred to Section 4.7.

Figure 7 shows the five tests of Kaolinite with heptane at a gradient of 300. In test 76 a water baseline was established, which was not done in the other tests. The variability of the test results is obvious, with two extremely low permeabilities established and three others varying over 1.7 orders of magnitude, which are well above the water baseline. The two tests on Kaolinite with heptane at a gradient of 100 (in one of which a water baseline was established) both behaved similarly to tests 44 and 76.

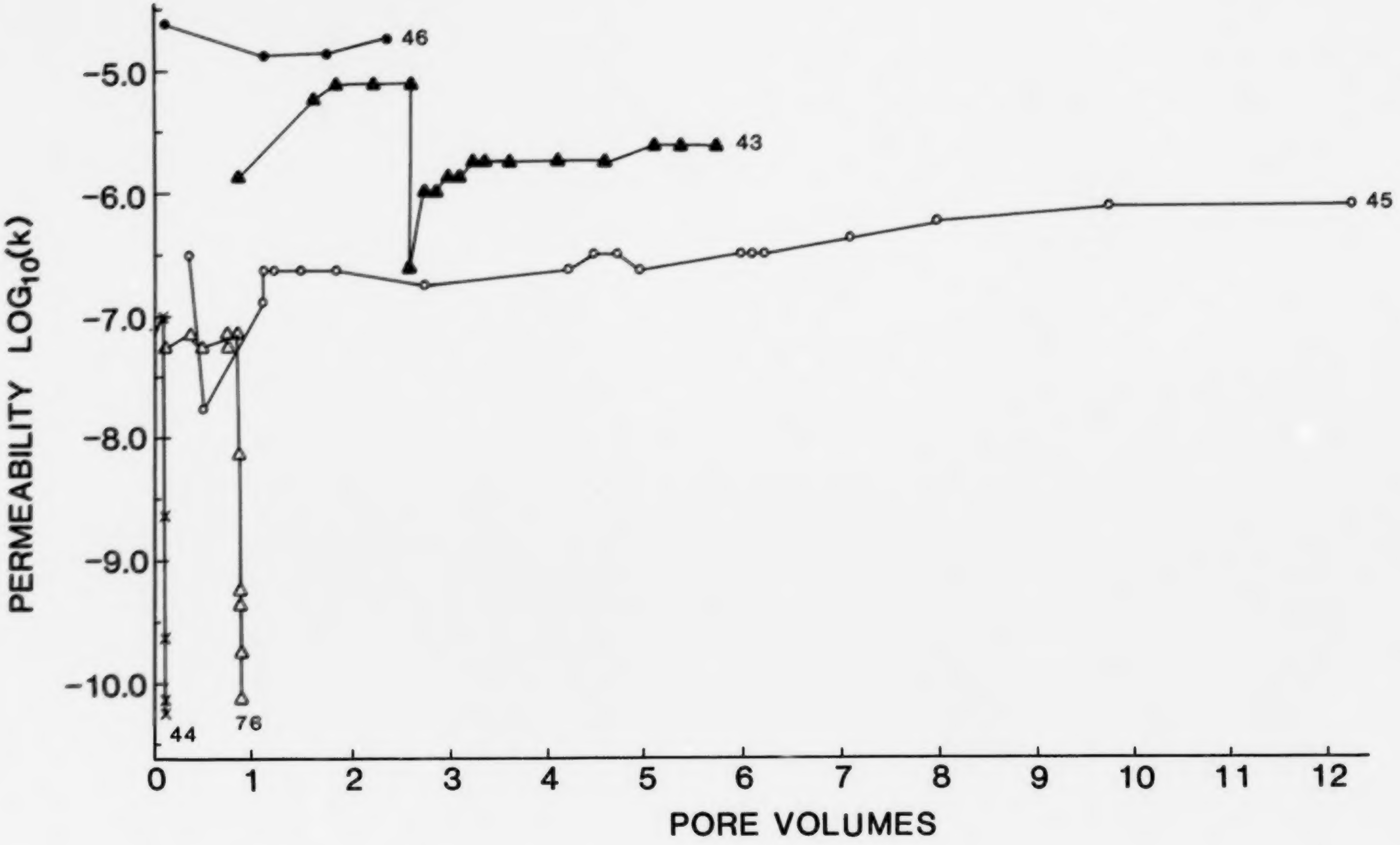
The inconsistency of the tests in Figure 7 is in sharp contrast to Daniel's tests on the same clay using either .01N $CaSO_4$ or methanol as the permeant. These tests were among the most consistent in the data set. This contrast would seem to indicate that the variability to be expected in tests with chemical permeants should depend strongly on the specific chemicals and clays involved.

Foreman and Daniel reported that in their early tests with heptane they had difficulties with the membrane encasement of the samples (20), which may be related to the variability shown in Figure 7.

Given that they had appeared to have achieved equilibrium, replicated tests on uniform clay soils which employed .01N $CaSO_4$ as permeant liquid and which were conducted under identical gradients and nominally equal molding water contents, generally gave estimates of the error standard deviation which were less than or equal to .15 on the logarithmic scale. Similarly, the estimated standard deviation resulting from Daniel's flexible-wall tests of methanol on Hoytville Illite and Lufkin Noncalcarious Smectite was .15 and on Kaolinite was considerably less (see Section 4.5). Five flexible-wall tests of Kaolinite with various concentrations of acetic acid were unreplicated, but appeared to be of lesser variability than the above benchmark would suggest.

To help interpret this standard deviation, suppose that on the log scale, replicated permeabilities were at least roughly normally distributed with a standard deviation of .15. Then about 95% of a large number of such replicates would fall within ±2 standard deviations of their average value. Replicated determinations would therefore typically fall within a band of ± 0.30 orders of magnitude. Since $10^{.30} = 2.0$, this would imply that generally such replicates would be no greater than twice the median permeability and no less than one-half this amount.

As examples in the following Sections will indicate, the amount of variability noted in some series of water tests in the data set was



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 7. Permeability of Kaolinite Clay with heptane (flexible-wall permeameters, i = 300)

considerably less than the rough benchmark described above would indicate. For the reasons noted in Section 3.3, however, it is believed that the interpretation of these perceived differences is often ambiguous and that attempts to generalize them may possibly lead to incorrect inferences. Furthermore, there are a number of other factors to consider which may lead to errors and variations which are considerably larger than this rough benchmark would suggest. Only two such factors are briefly introduced here, with more detailed discussions deferred to subsequent Sections.

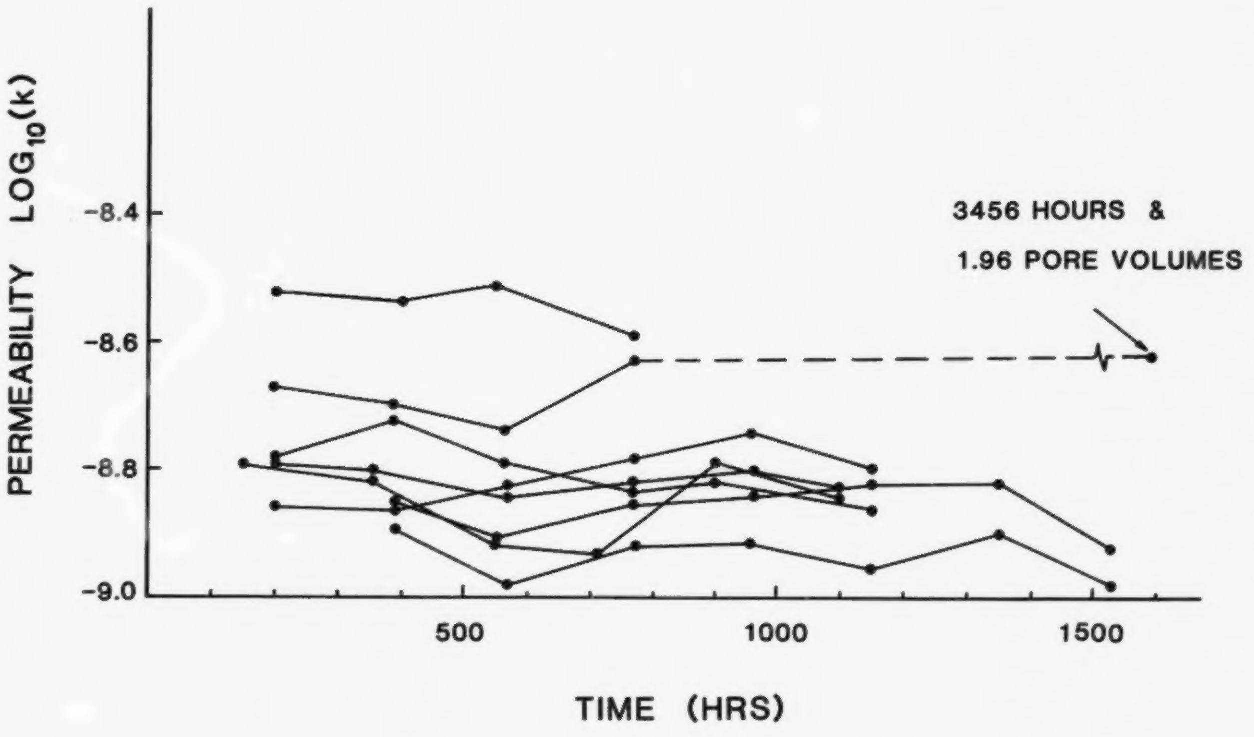
In a number of series of tests, even with water as the permeant, equilibration of tests required considerable time, both in terms of pore volumes and in terms of real time. In some such cases terminal permeabilities of longer-term tests systematically differed from their initial levels by amounts which were large relative to the between-test variability of replicated initial permeabilities. The bias in estimating the eventual equilibrium level which is introduced by premature termination of such tests would not be reflected in the variance of their measured permeabilities.

In the Brown and Anderson experiments of Figures 3 and 4, one test of each of the four clay soils was continued as a long-term control test to provide a comparison for the remaining seven cells into which chemical permeants were introduced. These were continued for between 1.89 and 2.66 pore volumes (1440-4968 hours). Shown in Figures 8-11 are the individual permeability measurements made on each of the 32 tests during the water permeation phase of the experiment, together with the permeability of each long-term water test at the conclusion of the experiment (these correspond only to the permeabilities marked with "X's" in Figures 3 and 4, since raw data on the other tests were not included in the available data).

Figures 8-10 show that the permeabilities of the Lufkin Noncalcareous Smectite, Calcareous Smectite and Kaolinite long-term tests are in reasonable agreement with the permeabilities of the shorter tests. In the case of the Hoytville Illite (see Figure 11), the long-term test increased about four tenths of an order of magnitude (151%) over its average level during the baseline period and about .67 orders of magnitude (367%) over the average level of all baseline tests. This test, together with other examples in which failure to quickly establish an equilibrium permeability was noted, is discussed in more detail in Section 4.4. The point of introducing the example here is to emphasize that the standard deviation of replicated short-term permeability tests may considerably overstate the actual precision of the tests, if in fact equilibration has not been achieved.

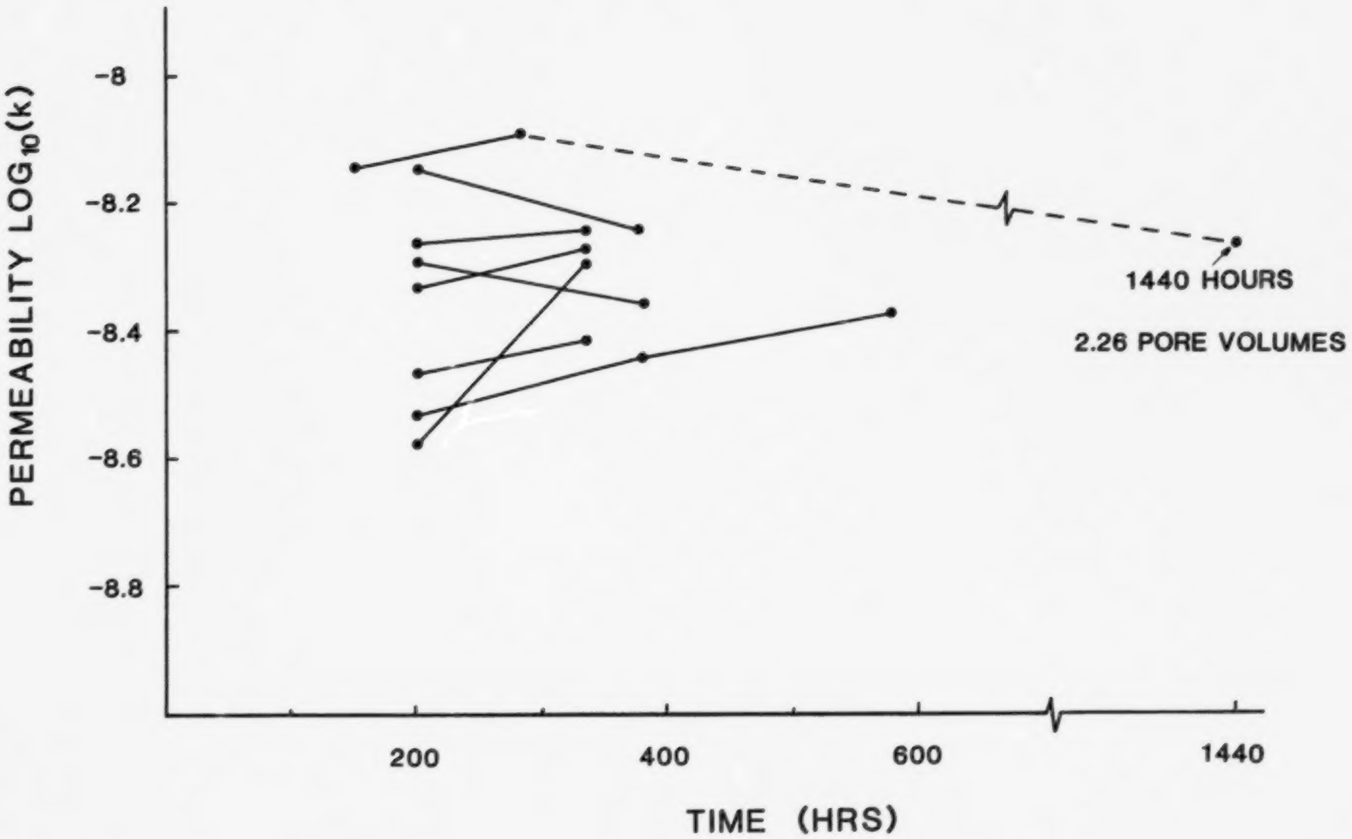
The level at which the tests discussed here have been replicated is not clear. In discussions with Brown, it was ascertained that the moisturization of the soils for individual specimens was generally done separately for each test, but that in some cases a number of specimens may have been compacted from a common batch of moisturized clay. Such uncertainty on the level of replication of tests was typically the case for the data which have been collected, with the exception of the data obtained from Peirce.

Replicated specimens at the within-batch level appear to be somewhat less variable than those replicated at the between-batch level, due not only to the presumably tighter control maintained on molding water content, but also due to enhanced homogeneity of the soils compacted into the cells and probably somewhat more uniform specimen preparation. This effect is conceivably



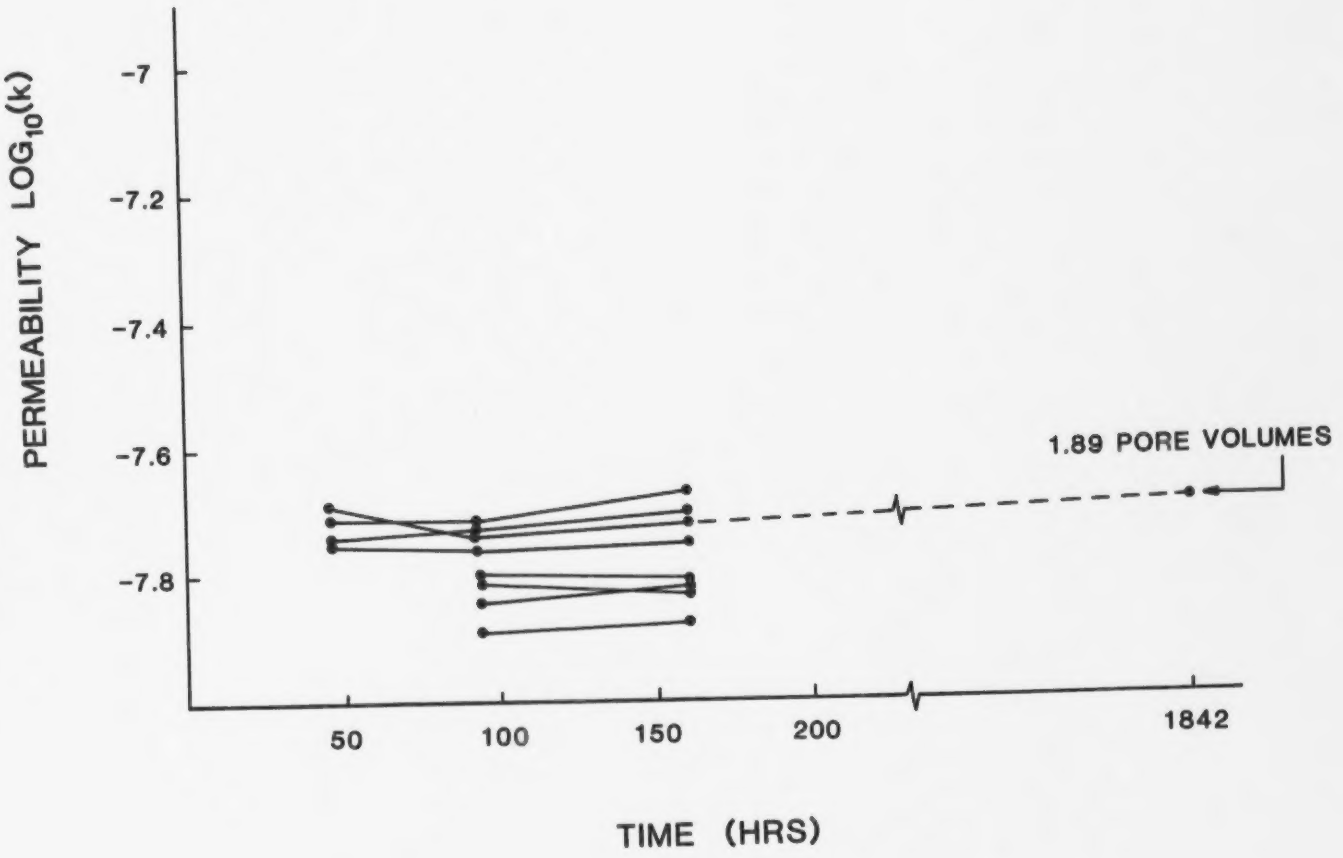
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 8. Permeability of Lufkin Noncalcareous Smectite with .01N $CaSO_4$ (rigid-wall permeameters)



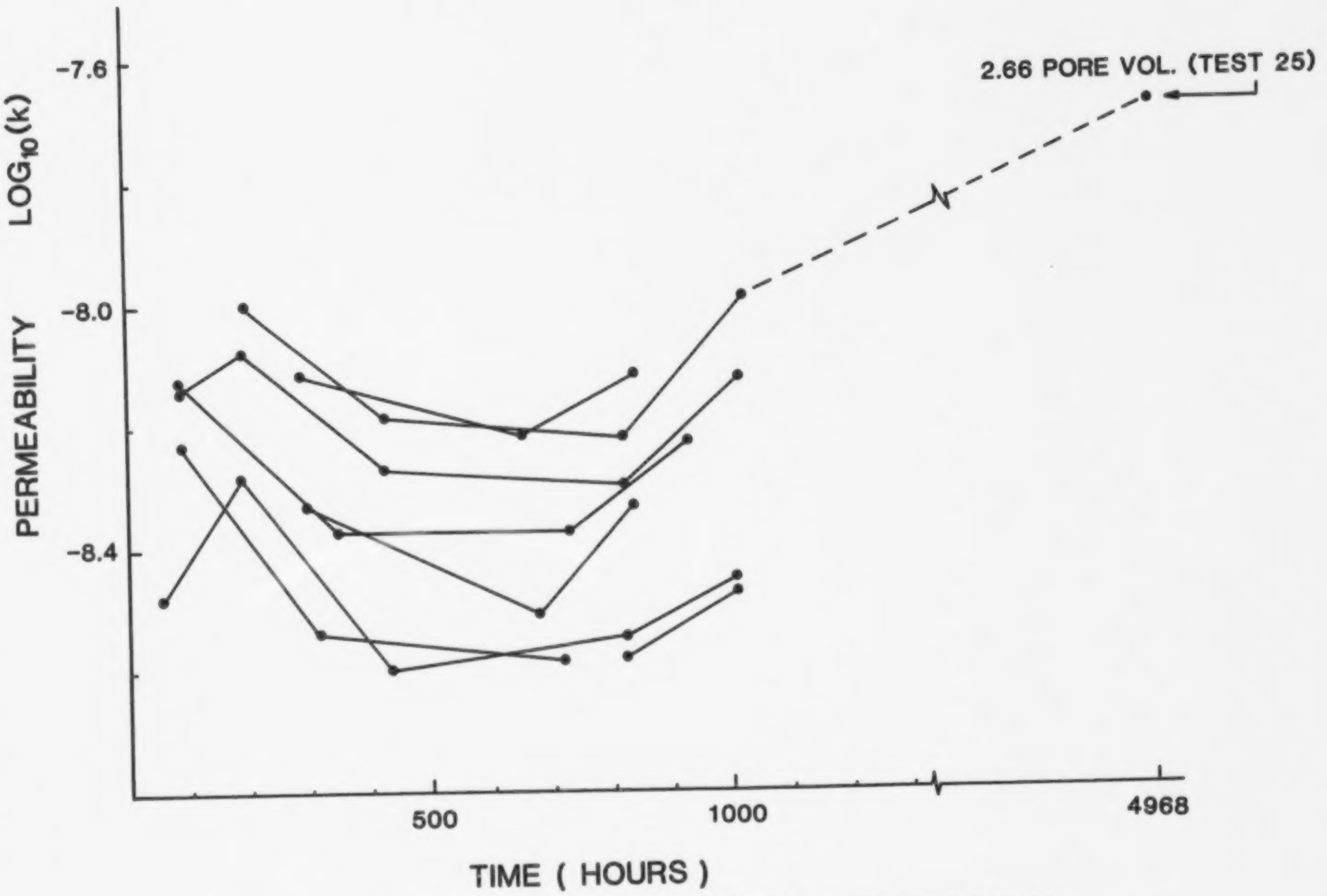
(NOTE: SEE APPENDIX A FOR A CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 9. Permeability of Calcareous Smectite with .01N $CaSO_4$ (rigid-wall permeameters)



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 10. Permeability of Kaolinite Clay with .01N $CaSO_4$ (rigid-wall permeameters)



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 11. Permeability of Hoytville Illite with .01N $CaSO_4$ (rigid-wall permeameters)

reflected in Figures 3 and 4, where some clustering of permeabilities may possibly be occurring. Such within-batch correlation of results may have the effect of understating the variability that may be expected to occur in replications which involve independent moisturization (at nominally controlled molding water content) for each test.

Similar remarks can be made for the correlation induced between tests run concurrently, due to the effects of environmental variables such as temperature. Although such effects are often relatively minor in comparison with the degree of between-test variability that may be observed, they are unmistakably present in tests which are known to have been run simultaneously (see Section 4.4).

4.2 LOCATION-TO-LOCATION VARIABILITY

The discussion of variability of permeability measurements in Section 4.1 concentrated on those components of variance having to do with specimen preparation and the actual permeability test itself. Although some variation due to soil heterogeneity is undoubtably present in test data from the research labs, its effect is to some degree controlled, since the clays used within each laboratory in the experiments were presumably relatively homogeneous. A separate component of variance which warrants attention is due not to variability of specimen preparation and testing, but rather to the heterogeneity of soils to the degree that might be expected to occur within a potential source of liner material. Some estimation of the magnitude of this source of error is needed, for example, to provide guidance on the amount of sampling required in order to obtain sufficiently precise estimates of such characteristics as the median permeability throughout the proposed source, or to perform statistical hypothesis tests concerning its suitability for use as a liner material.

None of the data available to these writers were originally intended to form the basis of a components of variance investigation. For this reason, there were no instances in the data set in which systematic replication was performed at both the between-location and within-location levels. However, data provided by private consulting firms did involve tests of samples taken from various locations within a number of sites. It was hoped that these data could be used to provide some quantification of the total variance due to all components, that is, between-location heterogeneity of soils and variation contributed during sample preparation and testing. If the magnitude of this total variability was substantially greater than that typically seen in research tests on uniform clay specimens, then a case could be made for attributing the difference to the location-to-location heterogeneity of the soils.

As will be discussed below, variation of test results obtained on specimens collected from different locations within sites did exhibit variability of a substantially greater magnitude than is generally found in research tests on uniform clays, even when sample differences with respect to the Unified Soil Classification System are controlled. However, while this observation is certainly consistent with the belief that soil heterogeneity is a dominant factor in the variability of permeability tests, there are a number of issues which complicate the interpretation of the data. These issues involve the question of test equilibration and also the degree to which statistical adjustment of the data, to account for differences in specimen preparation and test methods, is either feasible or logically proper.

To illustrate these comments, data which were provided by an independent soil testing firm will be discussed. They consist of 23 permeability tests of samples collected from four different sites, all of which employed rigid-wall permeameters and constant head methods. Water was used exclusively as the permeant. There was no replication of testing on soils from the same location. That is, all samples were taken from different bag samples within the four sites.

The number of pore volumes passed during the tests varied from less than .05 pore volume to more than one pore volume, with a median of .27 pore volumes. In terms of time on test, the longest was of about 121 hours in duration.

The amount of variability found among such tests should depend rather strongly upon the uniformity of the soil sources in question. Thus, in this analysis the writers considered the variability of permeability determinations taken from within soil sources which were geographically contiguous, each of which was confined to a single Unified Soil Classification, with Atterberg limits and grain size data restricted to relatively narrow bounds.

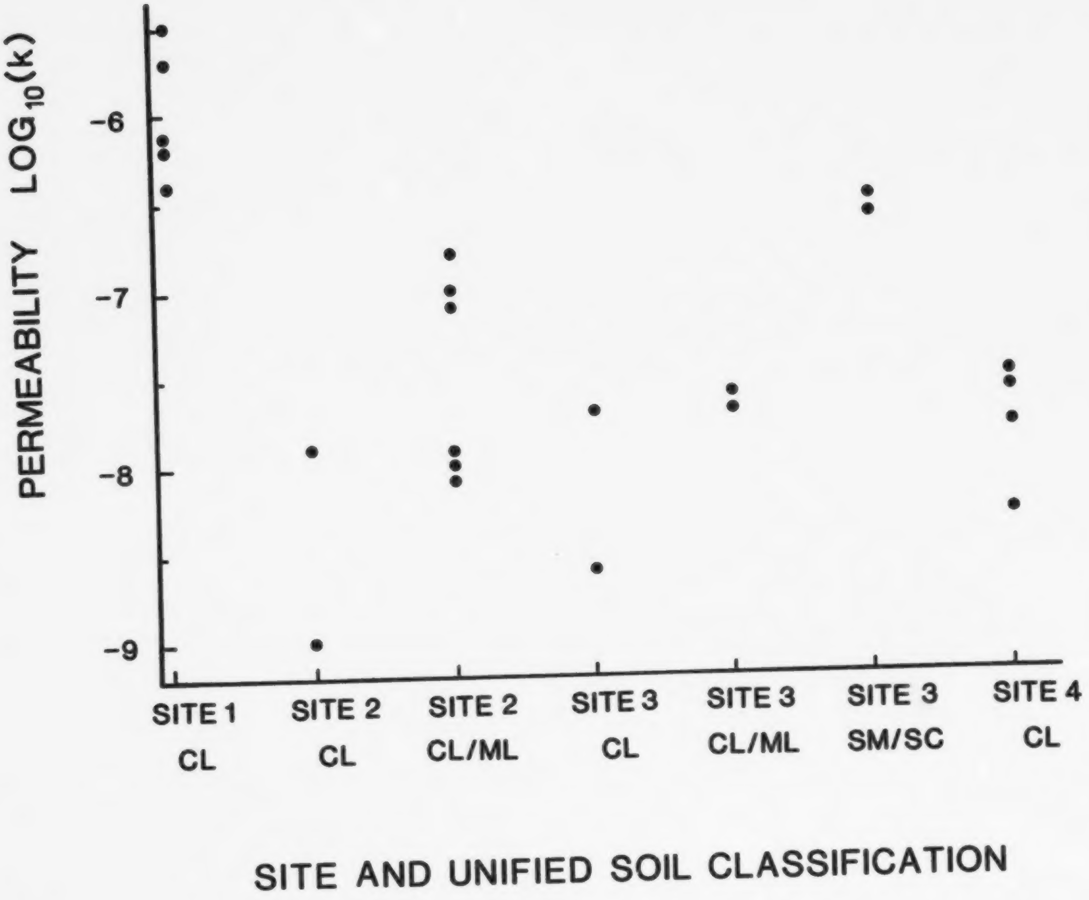
Within the four sites, samples may be grouped according to their Unified Soil Classification as shown in Table 3:

TABLE 3. SOIL GROUPS

SITE	UNIFIED SOIL CLASSIFICATION	NUMBER OF SPECIMENS IN GROUP	RANGE IN LIQUID LIMIT	RANGE IN PLASTICITY INDEX	RANGE IN PERCENT FINES	AVERAGE LOG10 PERMEABILITY
S1	CL	5	25-33	8-12	67-90	-5.97
S2	CL	2	30-32	15-17	64-72	-8.44
S2	CL-ML	6	18-20	5-6	52-58	-7.49
S3	CL	2	35-41	18-23	80-80	-8.14
S3	CL-ML	2	18-20	4-6	50-51	-7.64
S3	SM-SC	2	19-20	5-6	37-49	-6.55
S4	CL	4	24-33	10-17	65-68	-7.82

The results of the 23 permeability tests are plotted in Figure 12. Some noteworthy features of the plotted data are discussed below:

1) The variation of permeabilities within sites is extreme. Within either site 2 or 3, test results vary over nearly two and one-half orders of magnitude. Much of this variation is explained by differences in the Unified Classification of the soils



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 12. Permeabilities by site and Unified Soil Classification

2) Even restricting attention to the variability exhibited within site-Unified Classification groups, there are considerable differences. The CL-ML tests at site 2, for example, range over 1.2 orders of magnitude (1600%), while the two tests of CL clay at that site differ by more than an order of magnitude

3) In some cases, this within-group variability may be partially attributable to observable causes. For example, the Site 2 CL-ML tests appear to fall into two sub-groups whose permeabilities cluster around 10^{-7} cm/sec and 10^{-8} cm/sec, respectively. The initial void ratios at which these specimens were compacted varied from .23-.25 for the less permeable group and from .33-.40 for the more permeable specimens, even though their liquid limits, plasticity indices and grain size profiles are very similar, and the two groups of tests were on average compacted at nearly the same molding water content. Whether this is attributable to soil heterogeneity or differences in compactive effort is not clear

4) For purposes of comparison of the degree of variation in Figure 12 to that found in research tests, some numerical quantification is desirable. The variance of test results (on the log scale) pooled over the seven site-Unified Classification groups is 0.22, which is more than ten times as great as the variance observed in almost all of the water test series in the typical research setting using homogeneous clay soils and uniform test conditions. The corresponding standard deviation is $\sqrt{0.22} = .47$. In the presence of clustered data as in Site 2-CL-ML, however, the interpretation of these numbers is difficult.

If the entirety of the variability could be attributed to location-to-location heterogeneity of soils confined to particular sites and Unified Soil Classifications, then this component of variance would appear to dominate other causes of error. However, there are a number of reasons to question such a conclusion:

1) Most of the tests shown here are quite short relative to tests typically conducted in the research setting. In Section 4.4 it will be seen that equilibration of water tests may require considerable time, and it is possible that this factor could lead to variability of the tests in Figure 12 which is not related to soil heterogeneity. The majority of these tests involved only three permeability determinations over the course of the test, making an assessment of equilibration difficult or impossible

2) Tests within groups were not prepared to uniform specifications. In particular, there are meaningful within-group differences in molding water contents and initial void ratios. In theory, it is possible to adjust for such differences by means of analysis of covariance. The within-group ranges of variation of these parameters are small enough so that no obvious lack of fit is involved in linear adjustment of log permeabilities to the group average molding water contents and void ratios. If this is done, the apparent clustering of test

results in Figure 12 is eliminated, and the within group variance is reduced to 0.15 (standard deviation = .39). The effects of both molding water content and void ratio on permeability are statistically significant, at the .03 and .02 levels, respectively. After adjustment, the within-group variance is still qualitatively much larger than that found in research tests of uniform soils.

The covariance adjustment described above can be criticized on two grounds. First, it is known that the sensitivity of the permeability of soils to molding water content and void ratio differences varies from clay to clay. While there is no theoretical difficulty with constructing an analysis of covariance model to reflect this, the available data are simply not adequate to fit such a model.

A more interesting and important criticism pertains to the logical basis of the adjustment. For example, within-group differences in void ratio could be reflective of either soil heterogeneity or differences in specimen preparation. In the first case, it should not be adjusted for, and in the second case it should be. Likewise, much of the variability of molding water contents within groups was apparently due to the fact that specimens had different optimum molding water contents, and tests were generally run at molding water contents at or near this value. Thus adjustment for molding water content is also questionable if one is attempting to quantify the effect of soil heterogeneity on permeability.

3) Finally, certain other concomitant information is related to within-group variation of permeabilities, but the causal basis of this covariation is not clear. The best example of this is the relationship between permeability and the gradients at which the tests were run. Of all the available test information, gradient was the most strongly related to within-group variation of permeability. Data from research sources does indicate that the gradient at which a test is conducted may influence the nominal permeability of the test. On the other hand, a great deal of the observed relationship in this data set is probably due to the fact that the technicians who performed the tests chose gradients which were suitable to complete the tests in a reasonable amount of time, after an initial assessment of the permeability of the specimens on the basis of preliminary test results, visual inspection, and the initially observed rate of flow. Thus there are two reasons that explain the covariation between permeability and gradient, one that should be adjusted for (causal effect of gradient), and one that is spurious (choice of gradient following the initial assessment of permeability).

For these reasons, it becomes clear that the test results exhibited in Figure 12 should be taken at face value. If these data are reasonably typical, they indicate that the degree of variability to be found in tests taken from various locations within a site will likely be considerably greater than that observed in the research setting, even within soil sources which are reasonably uniform with respect to Atterberg limits and grain size information. Although this is consistent with the belief that soil heterogeneity is a primary factor in the explanation of this variability, any reasonable quantification of its effect is not likely to be obtained from an assessment of historical data such as these, but will require designed experimentation in which sampling is planned with this purpose in mind, replication is conducted at the various levels contributing to variability of permeabilities, and sample preparation and testing are carried out under uniform conditions.

4.3 EFFECT OF MOLDING WATER CONTENT ON PERMEABILITY

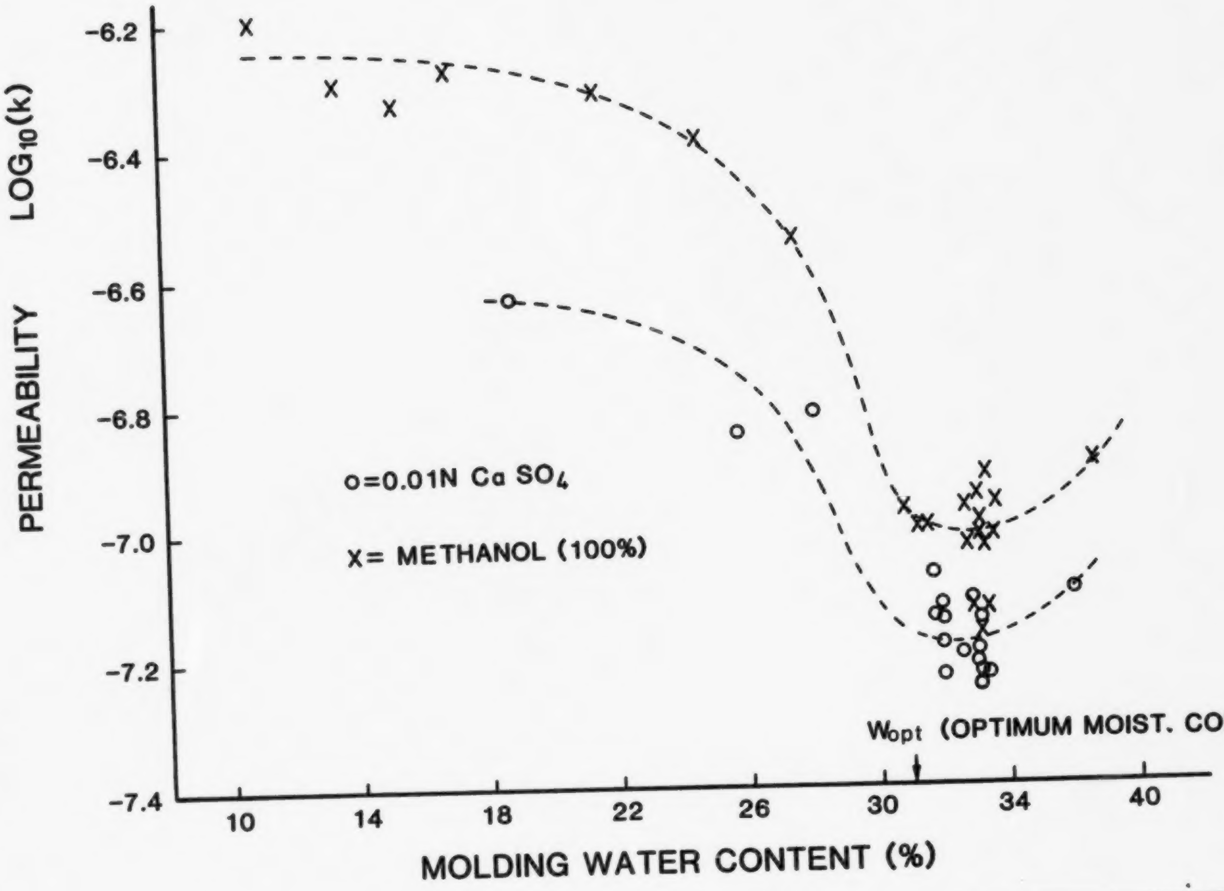
The permeability of a recompacted clay soil is very strongly associated with the water content at which it is compacted. Evidence of this relationship was found consistently in all the data sets investigated in which the molding water content was allowed to vary over an appreciable range.

A more or less typical shape of the permeability-water content relationship is illustrated by Figure 13, which is based on flexible-wall permeability tests by Daniel on Kaolinite, using both water (.01N $CaSO_4$) and 100% methanol as the permeant. At low water content, permeabilities are large and decrease rapidly as water content increases, up to a minimal point usually located several percentage points above the optimum water content (that is, the molding water content required to maximize dry density).

In the case of water, the strong influence of molding water content on permeability has been recognized for many years. Refer to Figure 14, which is from Mitchell (36) and which was cited by Roberts (44) to be the typical situation. Figure 13 is qualitatively identical to Figure 14. It also demonstrates that the relationship between permeability and molding water content in the case of methanol as the permeant has the same general characteristics as in the case where water is the permeant liquid. Unfortunately, this series of tests was the only one in the data set in which chemical permeant tests were conducted over a range of molding water contents wide enough to give a good indication of the relationship.

The magnitude of the effect of water content appears to vary for different clay soils. For example, the slope of the relationship in Figure 14 at its steepest point is roughly one half of an order of magnitude per one percent increase in water content, occurring at a water content of about 19%, or 1-1/2% above optimum. Since $10^{.50}=3.16$, this means that in this vicinity, a 1% decrease in molding water content would be expected to roughly triple the resulting permeability. In contrast, the maximum rate of change of log permeability for either methanol or water in Figure 13 is much smaller.

The data base contains examples of clay soils in which the sensitivity of permeability to water content is comparable to that found in Figure 14. Figure 15 shows log permeability versus molding water content for Calcareous Smectite permeated with water. These data come from rigid-wall tests of Brown and Anderson. Points marked with an 'X' are based on initial water permeation of eight specimens that was conducted prior to introducing chemical permeants. Points marked with an 'O' are the results of six preliminary tests with water as permeant which were not included in the data base but which were recorded in Tables A1 - A4 of their report (6).



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 13. Permeability of Kaolinite Clay with methanol and .01N CaSO$_4$ (flexible-wall permeameters)

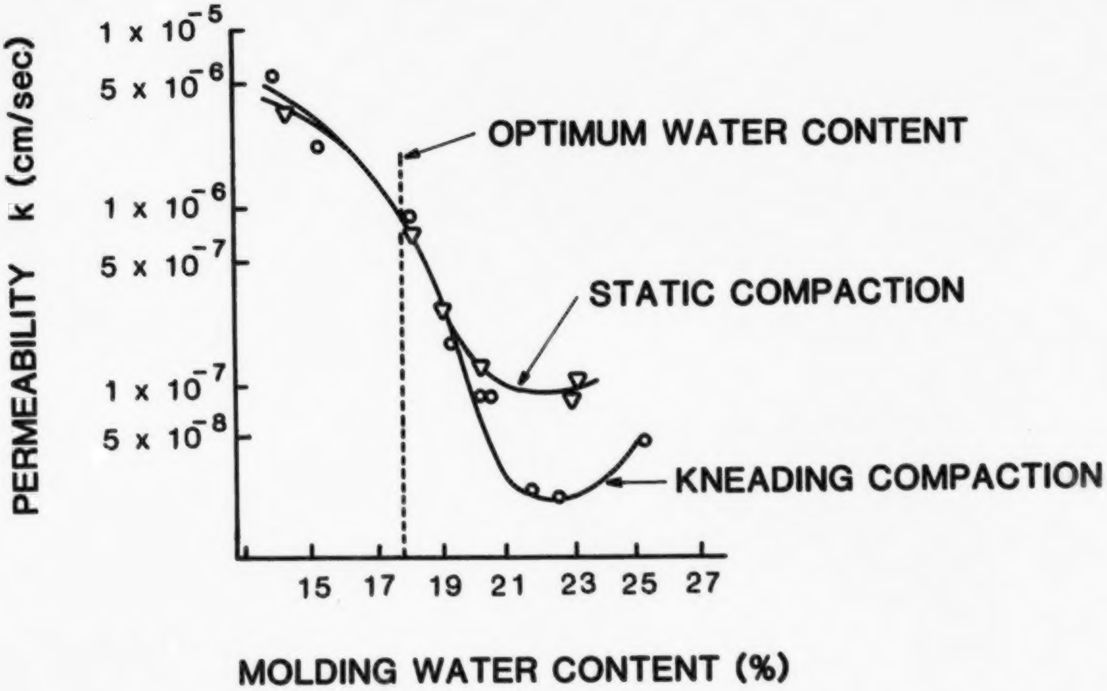


Figure 14. Permeability versus molding water content (Mitchell, 1976)

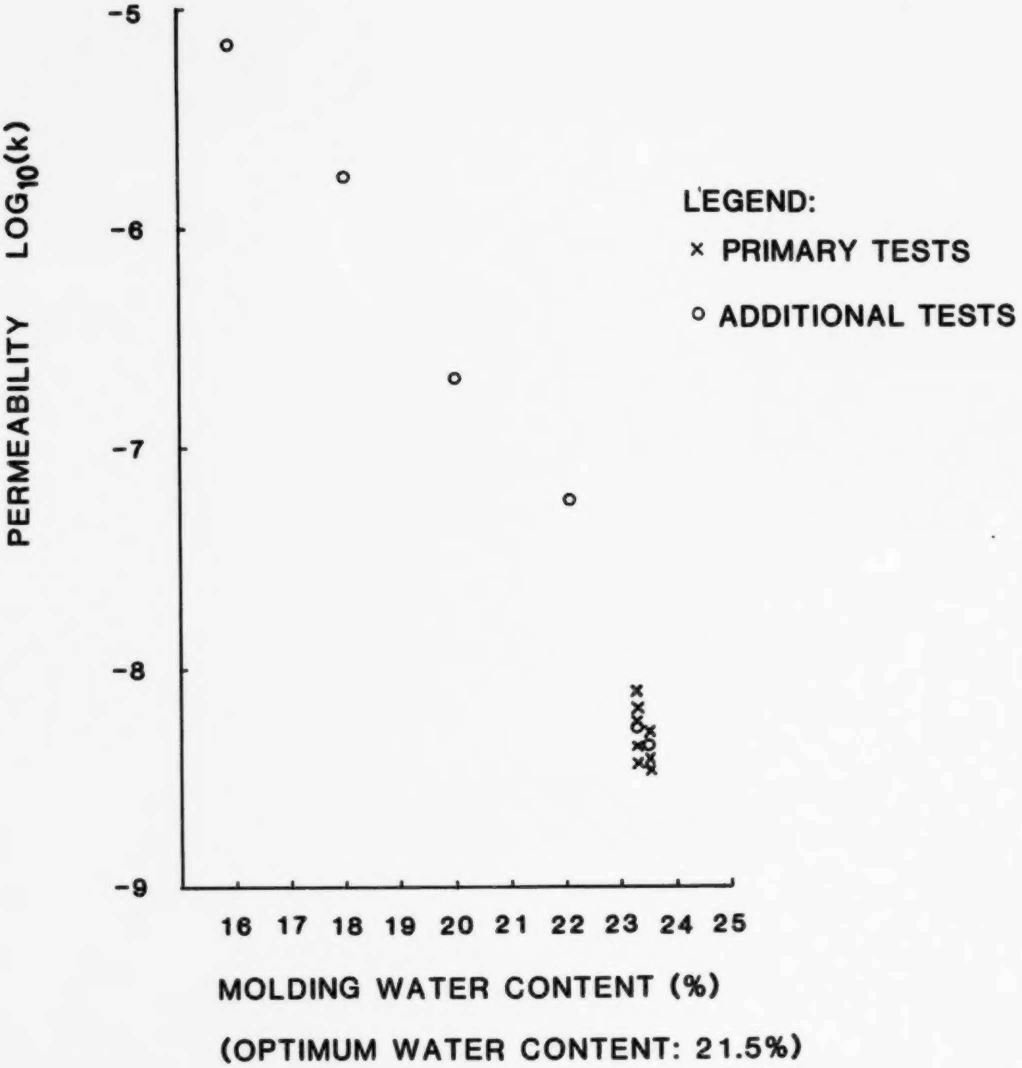
In Figure 15, the effect of water content is quite pronounced. Permeabilities decrease about three orders of magnitude as molding water content increases from 15.9% to 23.5%, or on average about 0.4 of an order of magnitude per percent increase in water content. Thus, in the general vicinity of optimum water content (21.5%), the sensitivity of permeability to molding water content for this clay soil appears comparable to that of Figure 14.

Figures 13-15 all indicate that the effect of the water content at which a clay soil is compacted is very significant in comparison to the variability of the permeability test itself, particularly in the general vicinity of optimum moisture content.

A potentially very important implication of the relationships shown in Figures 13-15 concerns the degree to which a laboratory permeability determination will be representative of the permeability characteristics that a proposed source of material would exhibit if it were to be used in the actual construction of a clay liner. To the authors' knowledge, little research has been undertaken which would permit an assessment of the degree to which water content and compactive effort may be controlled during liner application. If such parameters can not be closely controlled, then the sensitivity of permeability indicated in Figures 13-15 would imply that the results of laboratory determinations at the specified target values of these parameters may seriously misrepresent the future performance of the liner. This point is discussed further in Appendix B.10.

A second implication of the effect of molding water content on permeability is that it may inadvertently be confounded with the effect of treatments in comparative laboratory experiments. Consider Figure 16, which is a schematic representation of a situation in which such confounding could occur. In this figure, two treatments (e.g. comparison of flexible-wall and rigid-wall permeameters or tests run at two different gradients) are applied to each of ten tests. The resulting permeabilities would seem to indicate a difference between the two treatments when in fact the difference has been caused by average water content differences between those specimens to which the treatments were applied. If the water contents at which the twenty specimens were compacted were not precisely known, then it might be incorrectly inferred that the observed difference was caused by the treatments. Even in the case where the water contents were known, the effects of water content and treatments would be confounded; that is, it would be logically impossible to unambiguously separate their individual impact on permeability, so that no meaningful conclusions could be drawn from the experiment. Figure 17 shows an actual test series in the data base in which water content and possible gradient effect are confounded.

This problem of confounded effects is compounded by three factors. First, there is some doubt that water content determinations can be made accurately enough to recognize small differences whose effect may nevertheless be sufficient to create statistically significant, but spurious, results. These determinations are sometimes made by analyzing the clay skimmed off the top of the specimen and sometimes by preparing a parallel specimen for purposes of making the determination. In either case, the molding water content determination may not precisely duplicate conditions within the



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 15. Permeability of Calcareous Smectite vs. molding water content (rigid-wall permeameters)

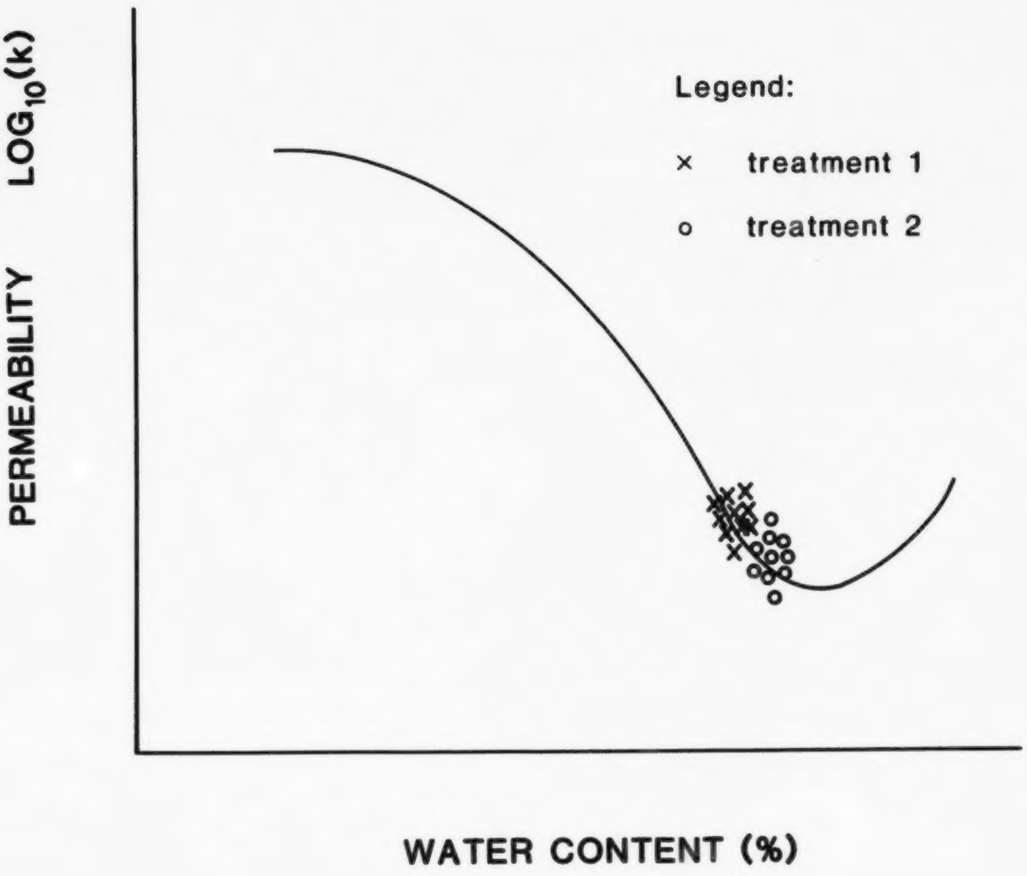
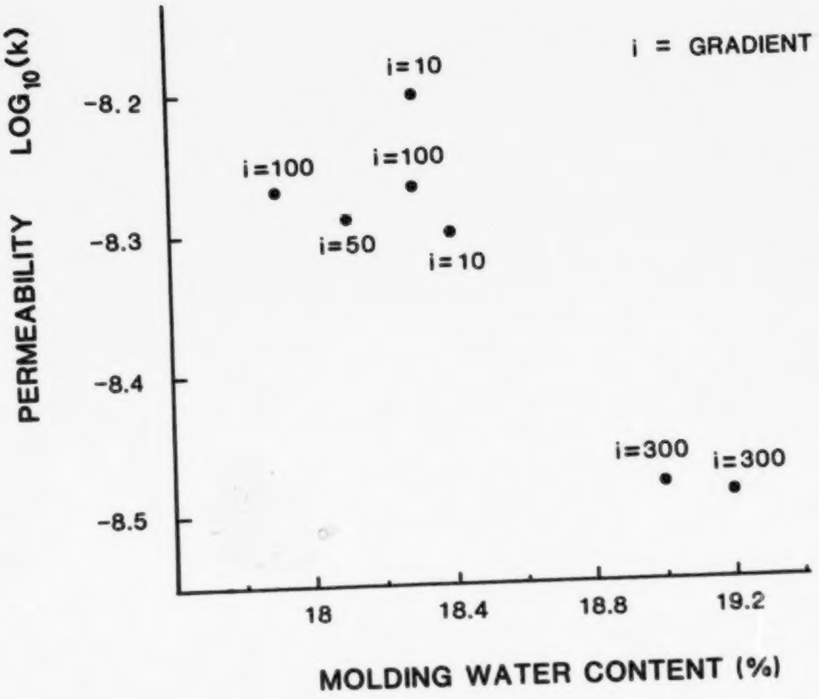


Figure 16. Confounding of treatment effects with molding water content



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 17. Confounded molding water content and gradient effects

permeameter itself. Secondly, there appeared to be few situations in the data base in which experimentation was consciously randomized with respect to the assignment of experimental material to treatments or to the order in which specimens were prepared. Finally, in some cases the preparation of specimens from batches moisturized at the same time was not systematically controlled, so that between batch differences in water content as well as differences in soil characteristics and compactive effort were neither controlled nor randomized.

Recognizing the strong effect of molding water content in the general vicinity of optimum, it was not uncommon for experimenters to conduct comparative tests at somewhat higher molding water contents, where the sensitivity of permeability is not as great. Even so, failure to randomize and to systematically control for batch effects may lead to spurious results due to other causes.

Methods of statistical experimental design which may be used to prevent extraneous factors from biasing treatment comparisons are discussed in Section 5.2.

4.4 EQUILIBRATION OF WATER PERMEABILITY TESTS

At the initiation of this data analysis project, it was anticipated that tests involving certain chemical permeants would be found that would not reach equilibrium of permeability over the allotted time, or would do so only after initial erratic behavior. Although this conjecture was often found to be correct, these writers were surprised to learn that the equilibration of tests using only distilled water or .01N $CaSO_4$ standard aqueous leachate were also much less likely to reach rapid equilibration than may commonly be supposed. In some cases, such tests were observed which equilibrated only after undergoing changes over considerable lengths of time (either in terms of real time or in pore volumes) which were large relative to the initial between-test variability of replicated specimens. In other cases tests were found to exhibit nearly monotonic increases or decreases over the entire test period. Still other tests evidenced extremely erratic behavior over the first several pore volumes.

Perhaps the most fascinating series of tests in the data base is from Peirce and consists of twelve rigid-wall tests of Faceville Clay prepared as part of a comparative study of the effects of chemical leachates. The soil required for the tests was moisturized to a molding water content of 23.7% (optimum moisture content was 23%) in a common batch, then compacted into rigid-wall permeameters and permeated with water (.01N $CaSO_4$) for 321.75 hours or, on avergage, 2.0 pore volumes. Then chemical leachates were introduced and the tests were continued.

Figure 18 shows the base 10 logarithm of the permeability of each sample over the last 100 hours of water permeation. These permeabilities were averaged over the eleven readings made on each sample during that period. The twelve log permeabilities range from about -7.7 to -7.5 (2.0×10^{-8} cm/sec $< k <$ 3.2×10^{-8} cm/sec) with an average of -7.59 (2.6×10^{-8} cm/sec and a standard deviation of 0.061.

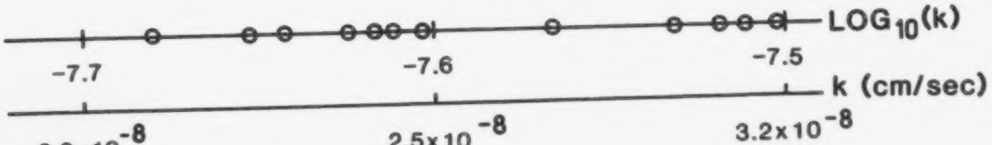


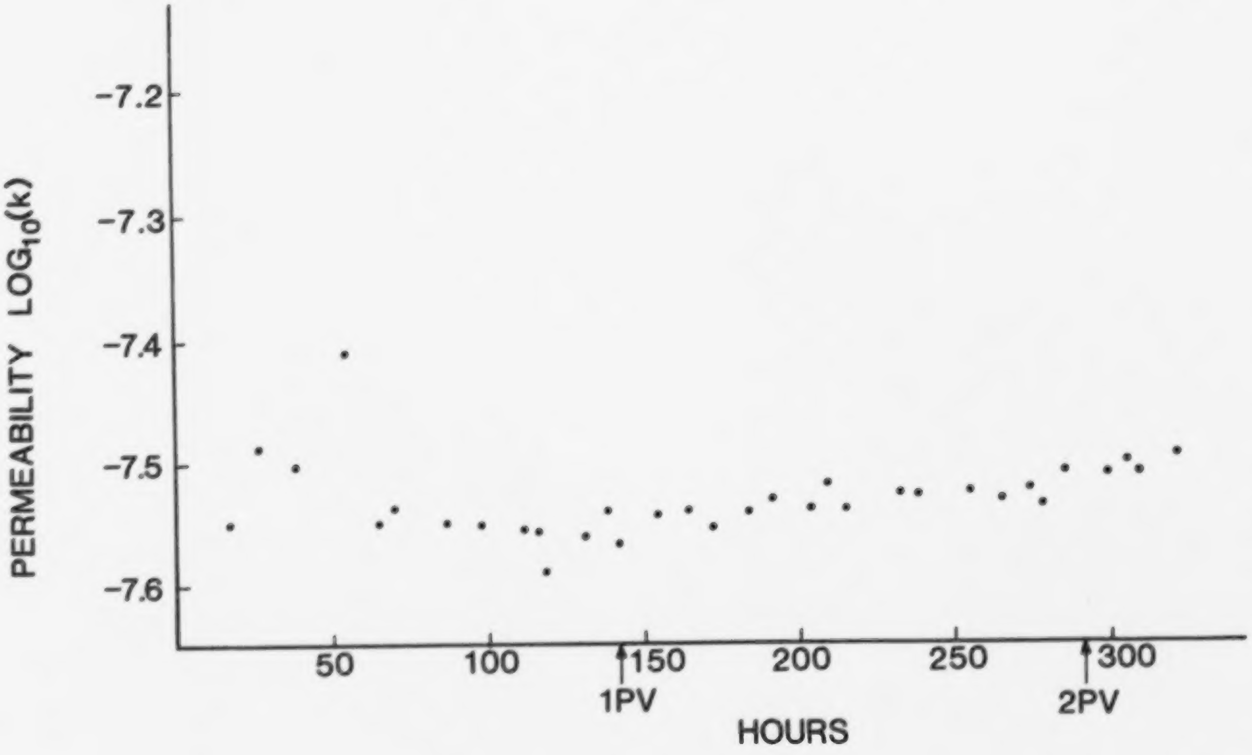
Figure 18. Permeabilities of Faceville Clay after 321.75 hours with .01N $CaSO_4$

Figure 19 shows the behavior of one of the twelve tests and is typical of the others. After brief instability, the permeabilities became fairly consistent at about 1/2 pore volume of flow. For the particular test shown, permeabilities appear to be increasing slightly during the last 100 hours of the test, but this effect seems quite small. A statistical test to detect systematic time trends across the twelve different tests is in fact significant ($p < 0.01$), but this formal result is practically meaningless, because the tests were run concurrently in the same lab and are jointly subject to minor changes in environmental conditions such as temperature and temporary fluctuations in the main pressure source. Furthermore, the estimated magnitude of the trend was only on the order of a 0.02 increase in log permeability, which could be attributed to such minor causes as a 2°C change in temperature. (Theoretically, the effort of a temperature change on permeabiilty should occur multiplicatively, through the effect of temperature on viscosity. For water as a permeant, the magnitude of this effect at usual laboratory temperatures is about .01 orders of magnitude per 1°C change in temperature.) From a practical point of view, the tests appear to be in equilibrium, as illustrated in Figure 20, which shows 100-hour average permeabilities for each of the twelve tests over the water permeation phase of testing. Based on the data reviewed in this project, had these tests been conducted by a geotechnical firm with the intention of determining the permeability of a soil, they would have been terminated much sooner than the 321.75 hours allowed by Peirce.

Figures 18 and 20 show that these within-batch replicated tests of water permeability exhibit excellent uniformity. The range of the twelve final permeabilities is, from Figure 18, only about 0.2 of an order of magnitude (or, since $10^{0.2}=1.58$, about 58%) which for this type of data is to be considered very precise. If, however, equilibrium is not achieved, additional error is contributed which may be large in comparison to this variability, as will be seen below.

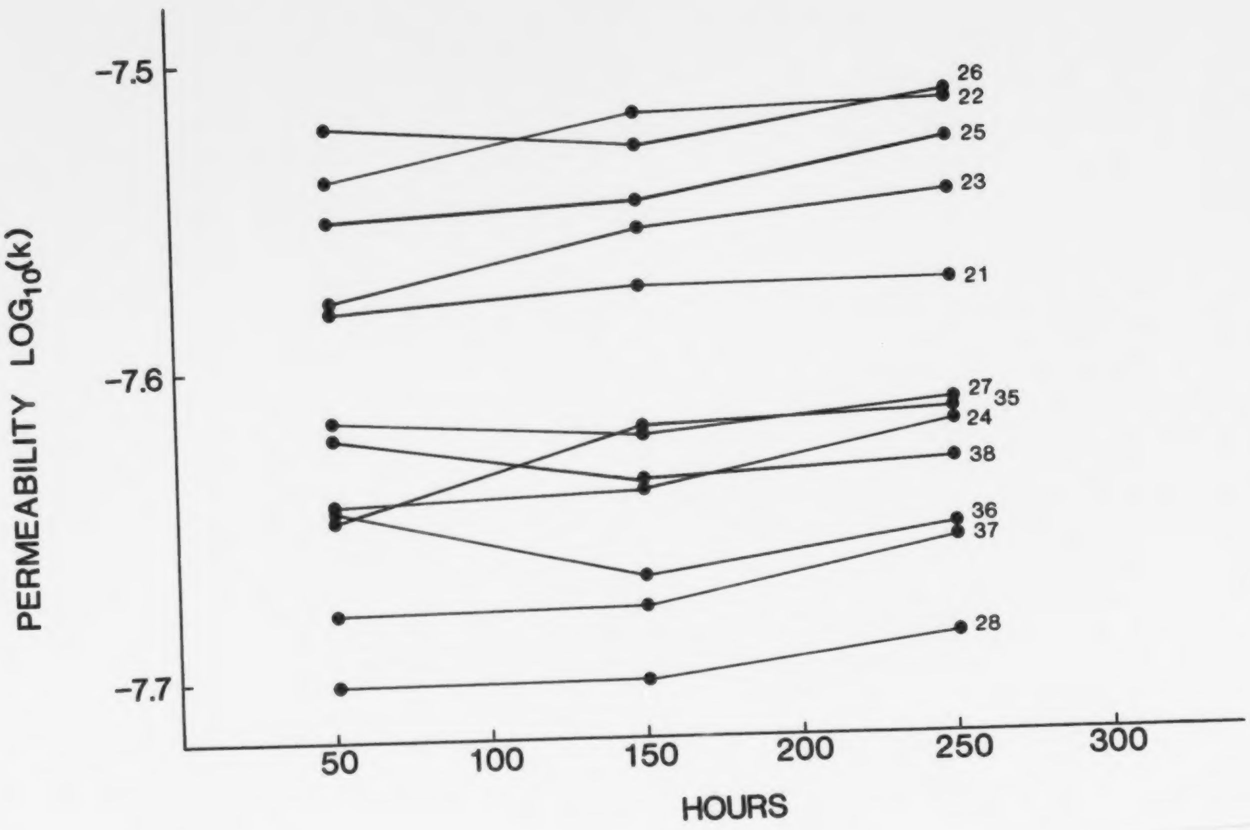
At the conclusion of 321.75 hours of water permeation, chemical leachates were added. Figure 21 shows the behavior of the tests. Each test is labeled with an identification number (21-28; 35-38). White dots at 271.75 hours represent the final water permeabilities analyzed above. White dots at 1020.9 hours represent the final permeabilities averaged over the last 100 hours of testing (970.9 to 1070.9 hours). Tests 36 and 38 were continued for 2391.33 hours. Black dots represent 100 hour average permeabilities. (Technically, the points plotted with black dots are medians rather than means, because occasionally an isolated permeability reading in the raw data was extremely high or low, having an undue influence on the average. This did not occur during the final intervals marked by white dots.)

The first impression given by Figure 21 is that addition of chemical leachates affected permeability, causing an average increase of about 0.44 orders of magnitude ($10^{.44}=2.75$). In fact, this is not the case; while four of the twelve samples received a dilute solution of $FeCl_3$ and four others received a $Ni(NO_3)_2$ solution, the remaining samples were used as controls with water continued as permeant, and these tests underwent permeability increases similar to the others. This becomes clearer from Figure 22, which shows the four water control tests in isolation, and from Figure 23, which shows the increase in log permeability for each of the twelve tests after the



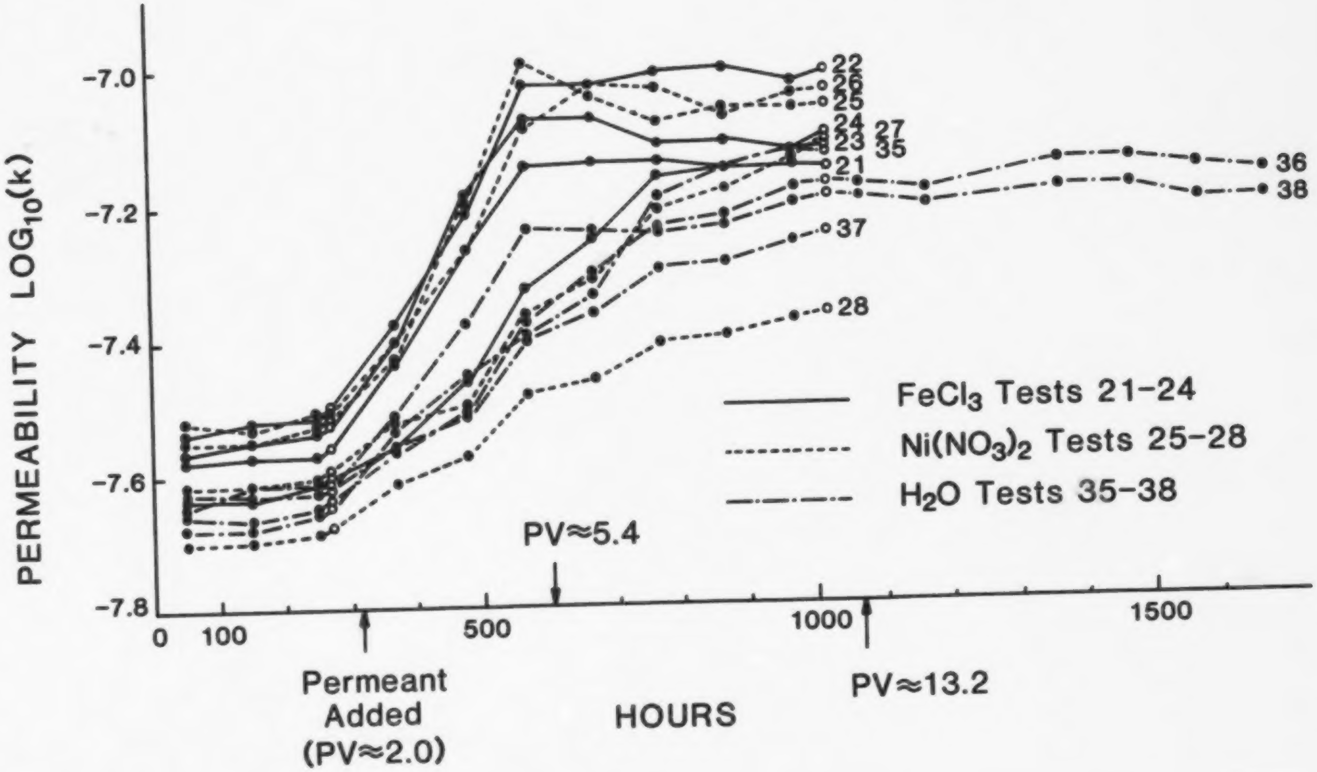
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 19. Permeability of Faceville Clay with .01N $CaSO_4$ (Test #25)



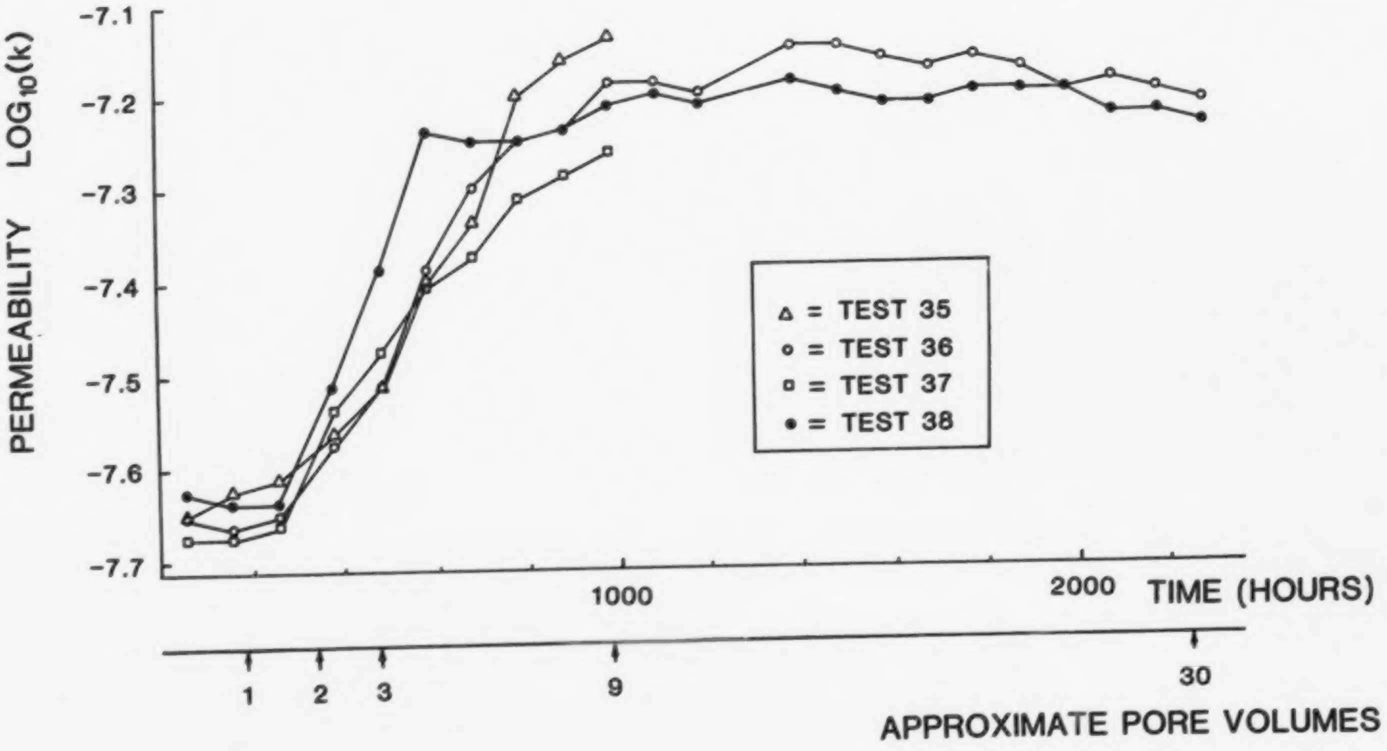
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 20. Initial permeability of Faceville Clay



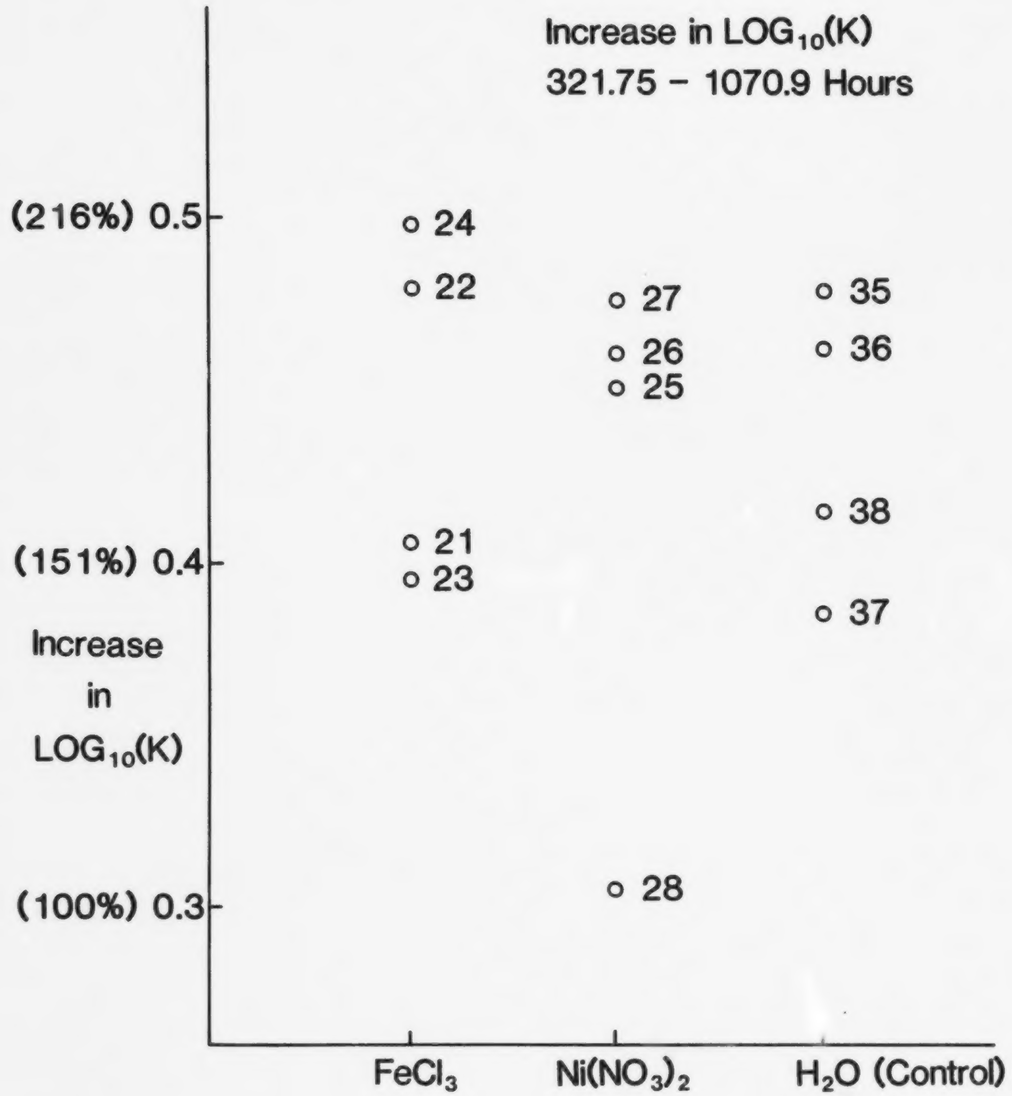
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 21. Permeability of Faceville Clay



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 22. Water control tests for Faceville Clay



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 23. Increase in $\log_{10}(K)$ of Faceville Clay between 321.75 and 1070.9 hours

introduction of chemical leachates. Statistical tests to detect an effect of the chemical leachates relative to water, which are based on the differences shown in Figure 23, do not indicate that the chemicals have increased permeability.

The cause of the observed increase is not known. A discussion of possibilities follows:

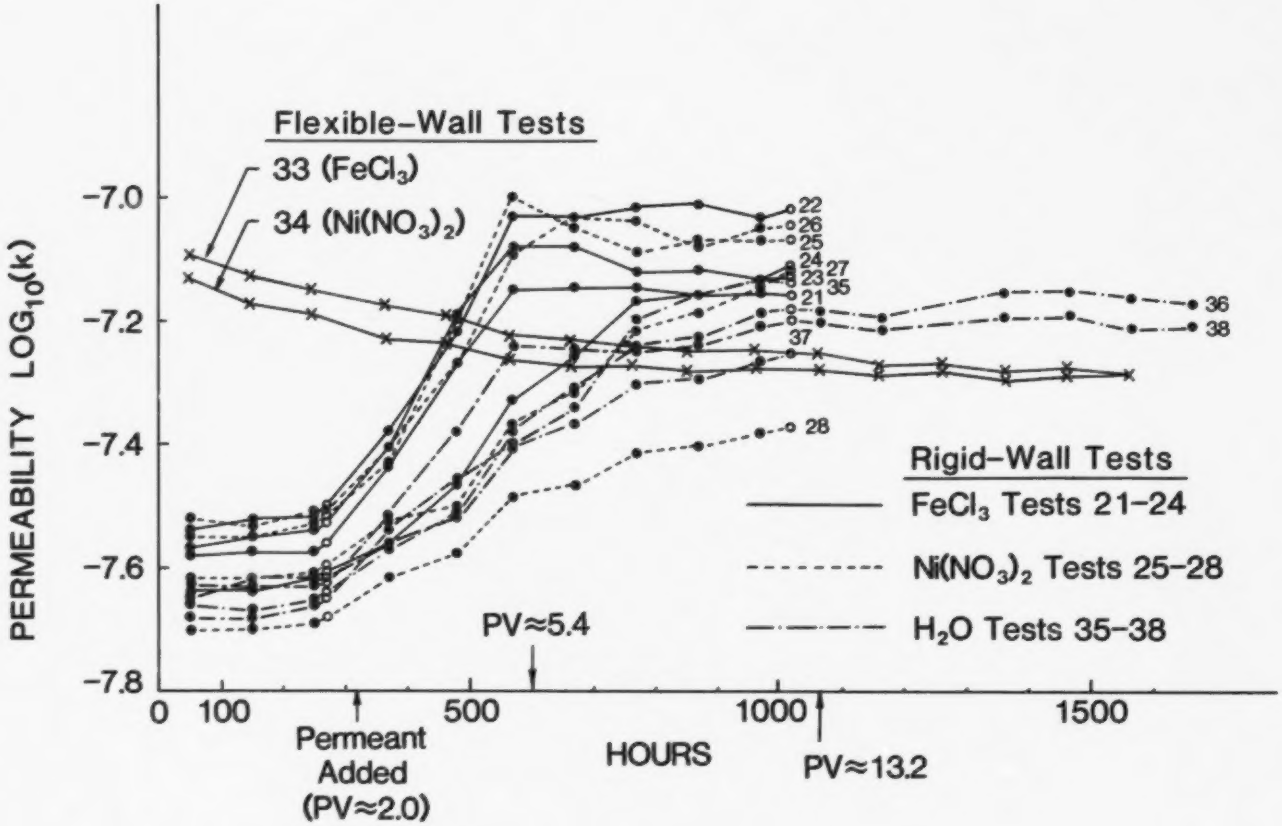
a. Change in laboratory environment: This possibility is indicated by the suprising degree to which all the tests behave similarly. There are certain environmental variables which are known to affect permeability and therefore would cause concurrently run tests to be correlated. Temperature, through its effect on the viscosity of water, affects permeability by about .01 order of magnitude per degree centigrade of temperature change (at reasonable temperature levels). An increase of the magnitude observed therefore cannot be explained by this mechanism. The tests were pressurized from a common source, although controlled by three separate regulators. However, a change on the order of that observed would require a tripling of net pressure over its nominal value which could not have happened. In fact, pressure gauges were recalibrated each day.

   Other tests (flexible-wall) were being conducted in the same lab during most of the period of increase. These did not exhibit any unusual behavior, which argues against attributing the change to some environmental cause.

b. Sidewall erosion: This is often cited as a cause for increasing permeability in rigid-wall permeameter tests and is a distinct possibility in this case. However, it seems hard to believe that all the tests would behave so uniformly if this were the mechanism at work. Additionally, the gradient used in these falling-head tests (approximately constant at about 150) is not excessive by the usual standards.

   Peirce also conducted two flexible-wall permeameter tests on Faceville Clay in addition to the rigid-wall tests discussed here, one using $FeCl_3$ as a permeant and the other $Ni(NO_3)_2$. Vacuum was applied to these two tests to remove air. As discussed in Section 4.6, the rigid-wall and flexible-wall tests were not intended to be used for a statistical comparison of relative permeameter effects, but an informal comparison is useful to this discussion. As shown in Figure 24, both flexible-wall tests start with permeabilities in the vicinity of $10^{-7.1}=7.9\text{x}10^{-8}$ cm/sec and decrease to virtually equal equilibria at about $10^{-7.29}=5.1\text{x}10^{-8}$ cm/sec, just slightly below the average level attained by the rigid-wall tests. If sidewall erosion were the cause of the increase of permeability of the rigid-wall tests, it might be expected that they would exhibit somewhat larger terminal permeabilities relative to the flexible-wall tests.

c. Sample disturbance caused by depressurization when permeant was switched from water to chemical: This also seems unlikely, because no switchover occurred in the control tests. Further, careful inspection of individual



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 24. Permeability of Faceville Clay
(Comparison of rigid-wall and flexible-wall permeameters)

graphs for each test indicate that perhaps the increase began slightly before switchover.

d. Time-dependent transition to a more flocculated state, associated with thixotropic hardening: Mitchell, et al. (35), have demonstrated that aging of certain clays prior to permeability testing may lead to increases in permeability consistent with the results seen in this study. It is possible that flocculation could occur during the test period in the rigid-wall tests, but was prevented in the flexible-wall tests due to the effective confining stresses.

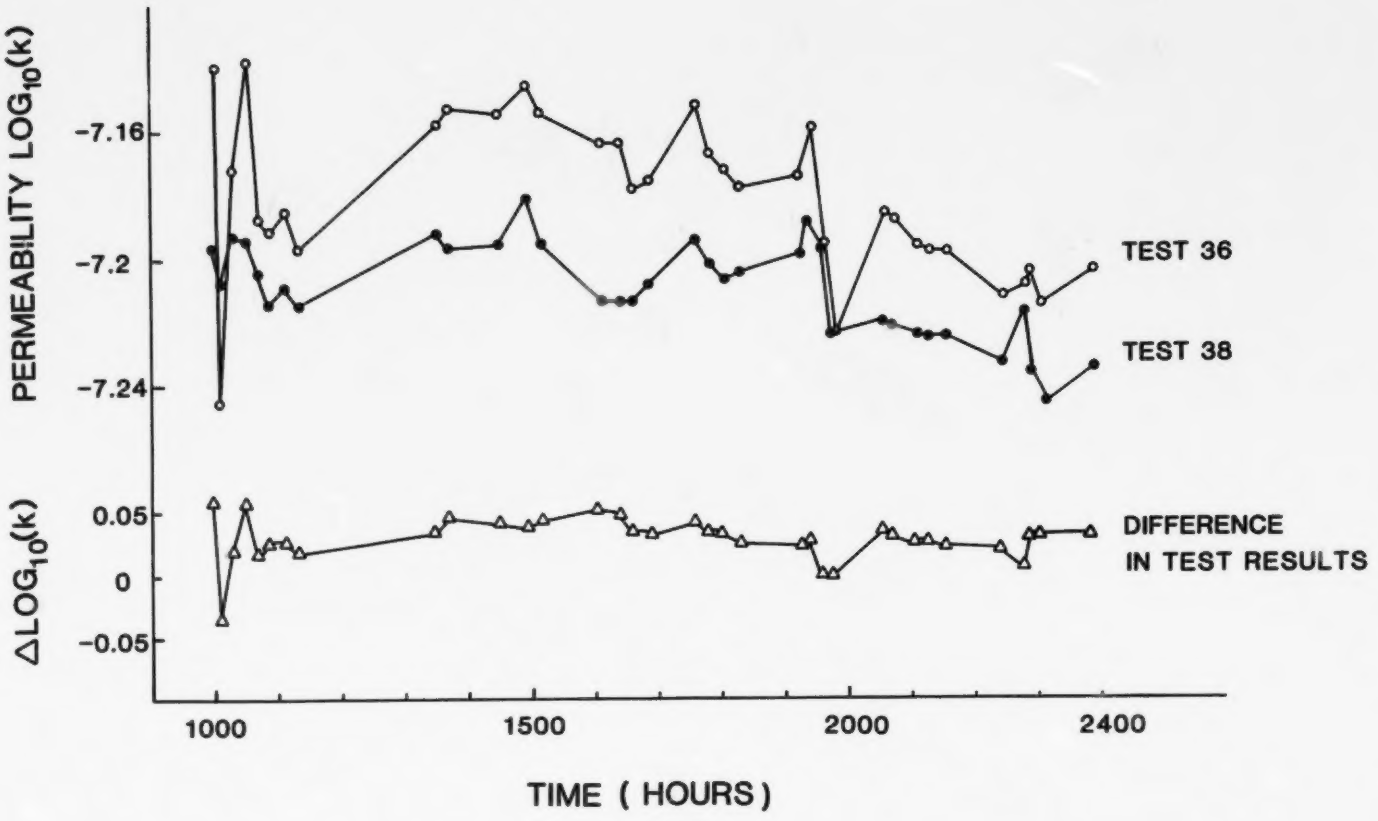
e. Slow approach to complete saturation, accompanied by increased permeability: This mechanism could conceivably explain the increase, together with the fact that the permeabilities of the vacuumed flexible-wall tests were considerably higher initially than those of the rigid-wall tests, while at equilibrium this was not the case.

f. A final possibility is that the .01N $CaSO_4$ has had a delayed effect on the permeability of the specimens. The control tests were permeated with .01N $CaSO_4$ throughout the experiment, while the inorganic leachates were in solutions with distilled water.

In the rigid-wall tests of Faceville Clay, the average log permeability increased from -7.59 to -7.15 (2.6×10^{-8} to 7.1×10^{-8} cm/sec), or over 7 times the initial standard deviation, even though the behavior of the tests was stable over the first two pore volumes. Even after as many as 10 pore volumes, the equilibrium of several of the tests is questionable. Furthermore, the variability of the determinations at 1070.9 hours is, from Figure 21, somewhat greater than at 321.75 hours, having a standard deviation of .097 (on the logamithmic scale) although this variability among tests may diminish somewhat as all tests reach equilibrium. This standard deviation gives a measure of the dispersion of permeability determinations which is relevant under the assumption that the introduction of permeants had little or no effect relative to the standard leachate. The within-permeants standard deviation of the determinations at 1070.9 hours is an almost identical .099 (on the logamithmic scale).

Water control tests 36 and 38 were continued over 2391.33 hours, accumulating roughly thirty-one pore volumes of flow. Figure 25 shows the behavior of the tests over the final 1400 hours. The points plotted in this graph are individual permeability measurements. The scale of the vertical axis is greatly enlarged.

The two tests are remarkably consistent over this period, with permeability in test 36 only about 7% greater than in test 38 (a difference of .03 on the log scale). The tests are seen to be noticeably correlated ($r = .73$). In Figure 25 it is reasonable to attribute this correlation to the common effect of environmental variables on the two tests, whereas prior to 1000 hours these effects were confounded with the equilibration process. The degree of correlation observed between the two tests is consistent with the hypothesis that underlying (and unobserved) environmental variables account for about three quarters of the within-test variance of log permeabilities.



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 25. Long-term test results on Faceville Clay with .01N $CaSO_4$

Over the 1400 hours shown in Figure 25 the permeabilities of both tests appear to drift over a range of about .05 orders of magnitude, or roughly an 11% change. This drift is of a magnitude which could be explained by known environmental causes. A .05 order of magnitude drift could be attributed to a 5°C change in average temperature and its resulting influence on the viscosity of the permeant, or a drift in net presssure of about 1 psi from its nominal value, or a combination of both factors in lesser degree. Although exact data on lab temperature was not available over the time period in question, it was noted that air conditioning was installed at some point during the experiment. Speculatively, this could account for the very slight decrease in permeabilities over the latter portion of the test.

The remarkable feature of Peirce's Faceville series of tests is not that they increased by nearly one-half order of magnitude prior to equilibrating; such behavior was observed in a number of tests using distilled water or standard aqueous leachate. What is most interesting is that they did so after exhibiting extremely stable initial behavior, both in terms of the smoothness of recorded permeabilities within the tests (Figure 19), and the consistency of permeability levels across the twelve tests (Figure 20).

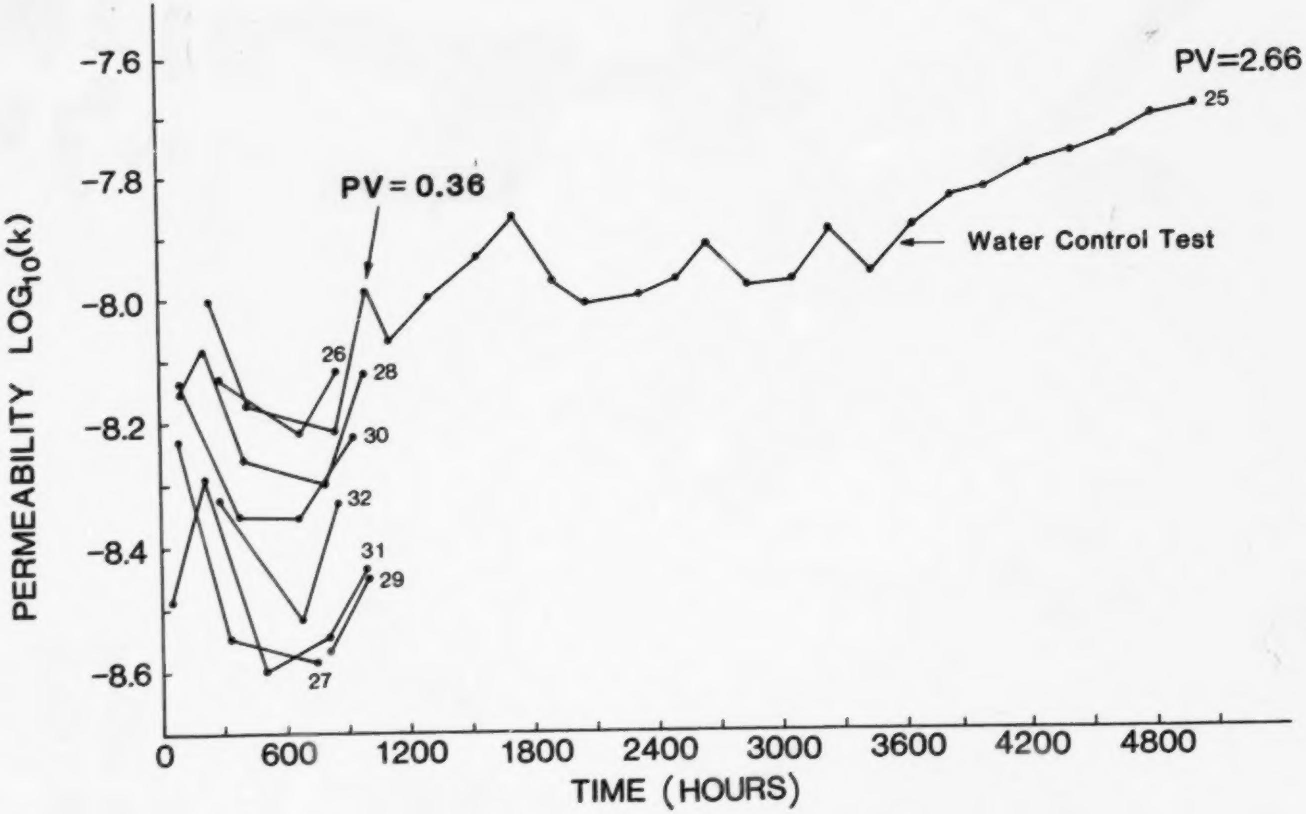
Figure 26 shows the water control test of Hoytville Illite done by Brown and Anderson, which was discussed briefly in Section 4.1.

In their report (6), Brown and Anderson discussed the behavior of this test. They noted that its increase in permeability over the course of the experiment was minor. For their purposes this was clearly true, since the permeability changes caused by the introduction of chemicals to the other cells in their experiment generally were extremely large relative to that experienced by the control test. However, from the viewpoint of this discussion the increase in permeability was not insignificant. At the termination of the test, its permeability is $10^{-7.69}=2\text{x}10^{-8}$ cm/sec, an increase of about 0.4 orders of magnitude ($10^{.4}=2.51$, an increase of 151%) over the permeability of the test based on its total flow during the baseline period. This increase is not small relative to the standard deviation of log permeabilities of the baseline tests, which was 0.15. Furthermore, Figure 26 gives no indication that the test has equilibrated even after more than 200 days (2.66 pore volumes) of testing.

Figure 27 shows four rigid-wall tests on Elgin Kaolinite Clay, run at a gradient of 91 with .01N $CaSO_4$. Test 39 falls about one order of magnitude from its initial permeability over the first .1 pore volume until it stabilizes after about 550 hours. Test 10 shows very erratic initial behavior before equilibrating about one-half order of magnitude below its initial permeability.

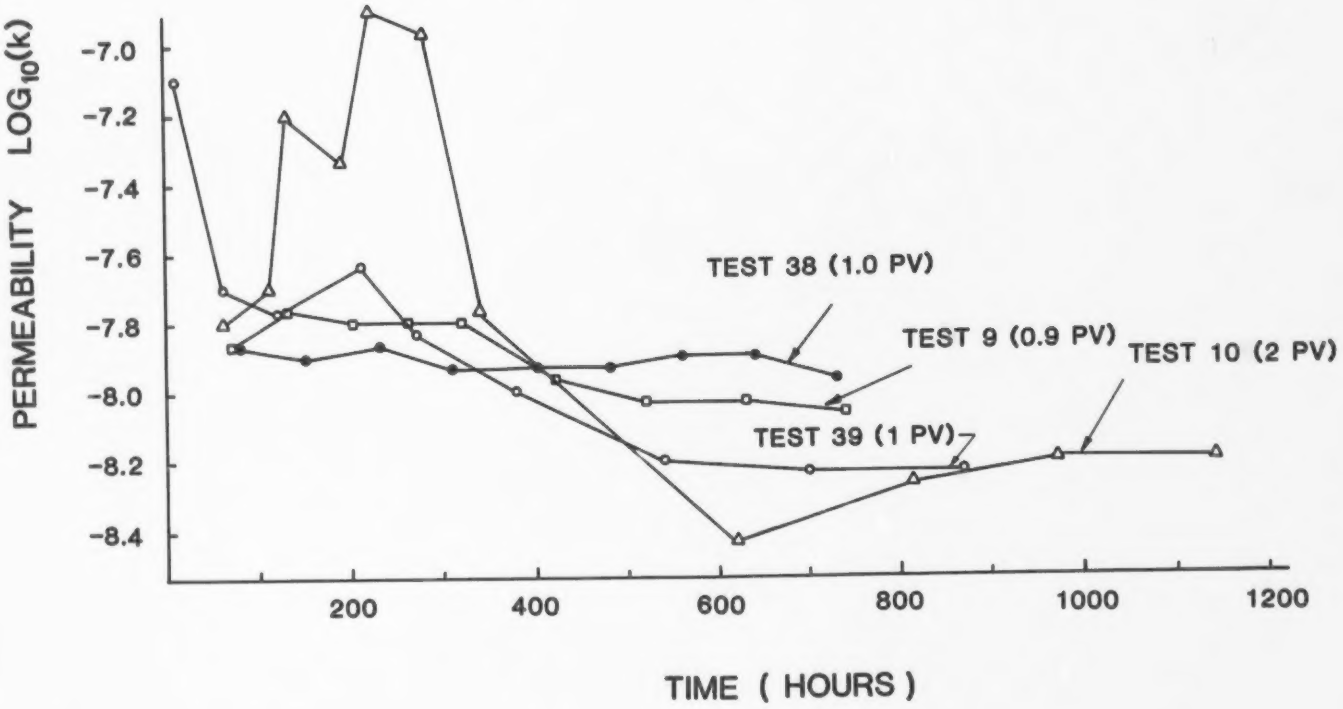
Figure 28 shows .01N $CaSO_4$ tests of the same clay at a gradient of 181. Test 41 displays an order of magnitude jump in permeability after having passed .9 pore volumes.

Figure 29 shows the third longest water test (in terms of pore volumes) in the data set. It employs distilled water and a rigid-wall permeameter on a second Kaolinite Clay, and was performed by Daniel. It again demonstrates that permeabilities over the first several pore volumes may be erratic. Here, an



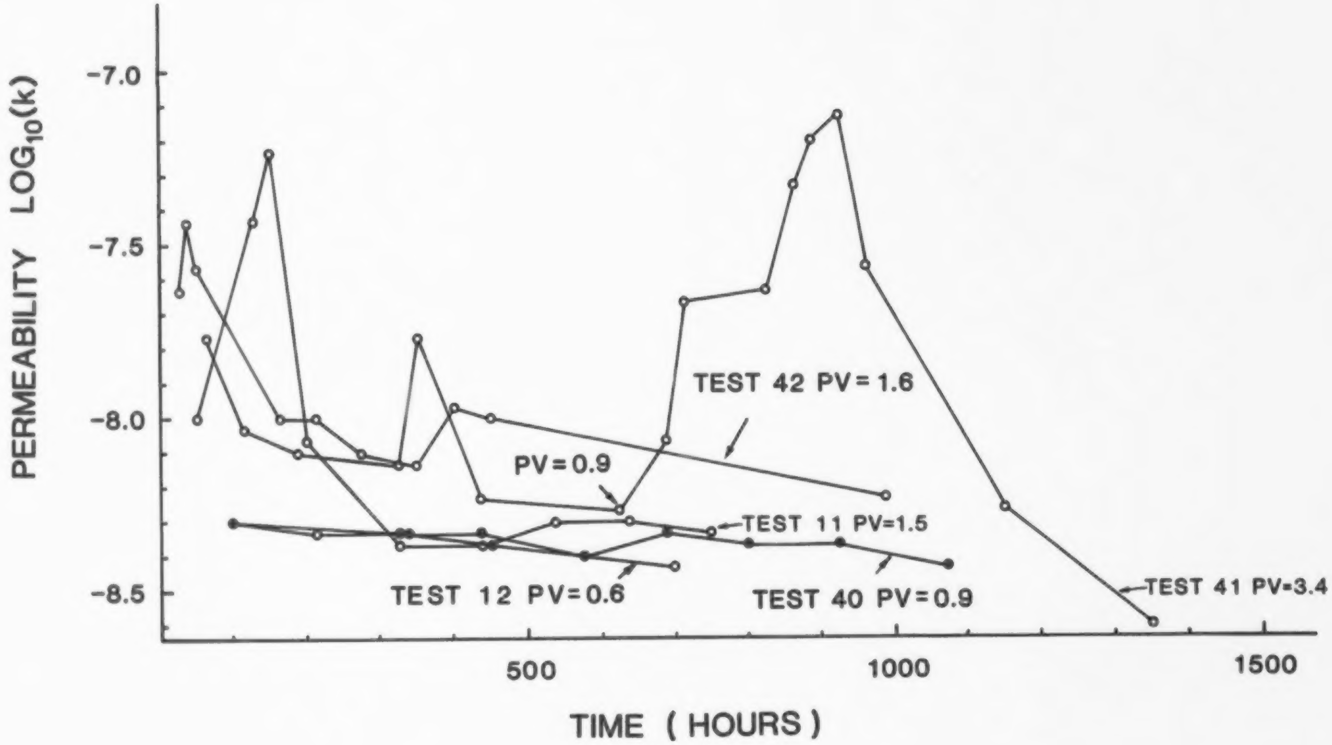
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 26. Water control test of Hoytville Illite with .01N $CaSO_4$ (rigid-wall permeameters)



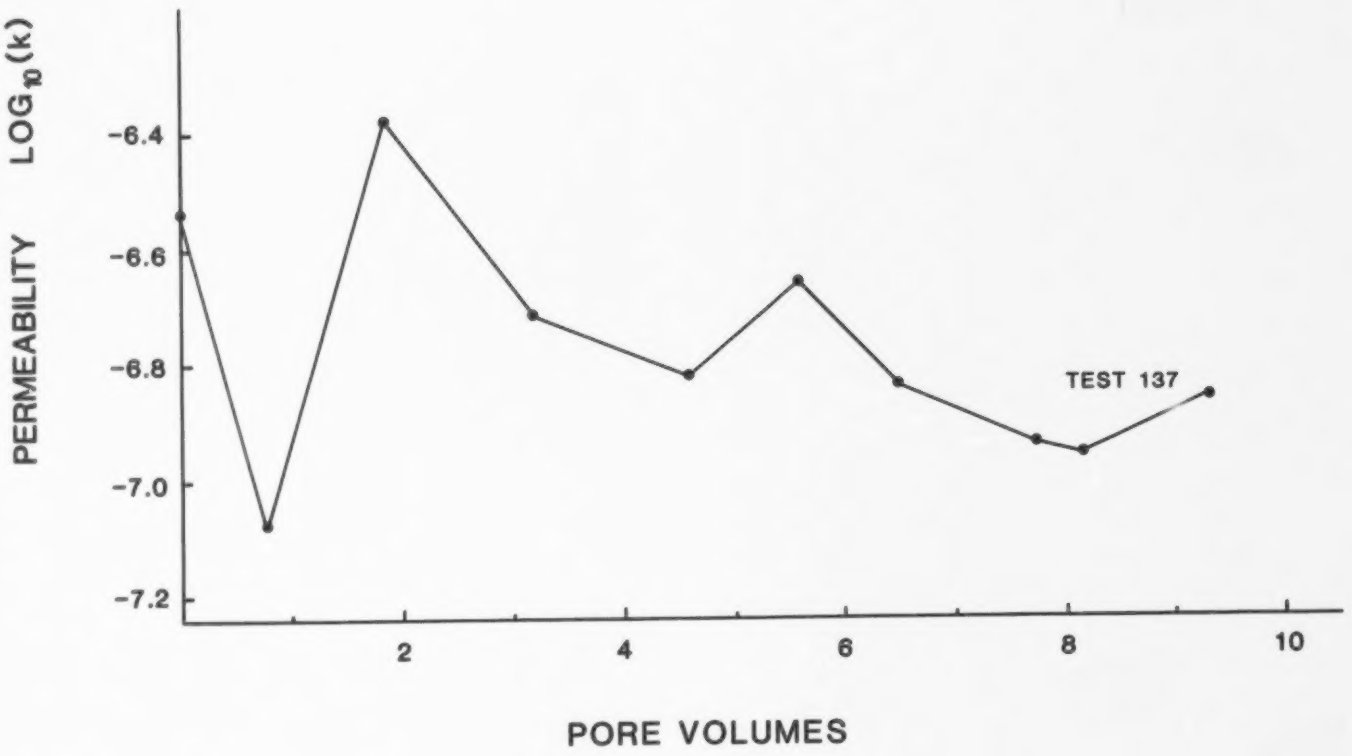
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 27. Permeability of Elgin Kaolinite with .01N $CaSO_4$ (rigid-wall tests with i = 91)



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 28. Permeability of Elgin Kaolinite with .01N $CaSO_4$ (rigid-wall test with i = 181)



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

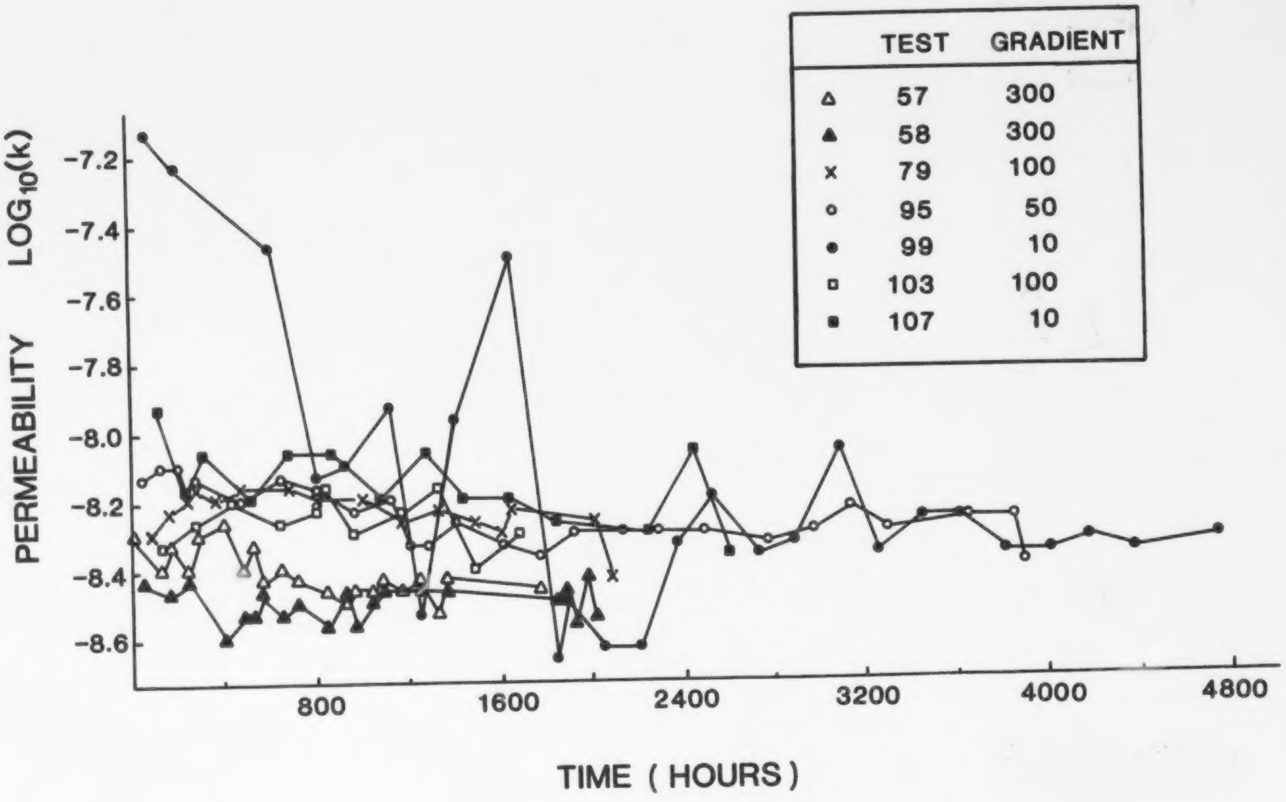
Figure 29. Permeability of Kaolinite Clay with distilled water (rigid-wall permeameter)

initial determination based on the first .04 pore volumes is $10^{-6.53} = 2.9 \times 10^{-7}$ cm/sec. Permeability first decreases and then increases, with the average rate of flow over the second pore volume about 5 times (.7 of an order of magnitude) greater than over the first. By the termination of the test, stability appears to have been achieved, with the average rate of flow varying by about .1 order of magnitude over the last four pore volumes. At the given conductivity, this test should have required about 20 days to complete.

In the discussion of flexible-wall and rigid-wall permeameters of Section 4.6, it is indicated that in a limited number of comparisons under matched conditions, flexible-wall permeameters seemed to be somewhat more inclined to achieve a stable equilibrium permeability than rigid-wall permeameters, at least for chemical permeants. However, this does not imply that considerable changes in permeability level will not be experienced prior to achieving equilibrium, even with .01N $CaSO_4$ as permeant. Figure 30 shows a series of seven triaxial tests of Hoytville Illite conducted by Daniel using 0.01N $CaSO_4$ at gradients ranging from 10 to 300 and at molding water contents ranging from 17.9% to 19%. Of particular interest is test 99, which falls over an order of magnitude prior to reaching equilibrium and which shows very erratic behavior over the first half-pore volume. At the reduced gradient of 10 at which this test was run, it appears to require about 100 days of testing before this specimen stabilized. In the same way, the permeabilities of a number of Daniel's flexible-wall tests of Lufkin Noncalcareous Smectite with 0.01N $CaSO_4$ fell by about an order of magnitude over the test period and did not completely equilibrate (see Figure 38 in Section 4.5). Most of these tests were extremely short in terms of total pore volumes of flow (typically .06 -.30 pore volume) but still, at the gradients used in the experiment even those tests which did not equilibrate would have been run for over one month.

The purpose of discussing these examples is not to imply that the behavior of permeability tests is always erratic or involves substantial changes in permeability level before equilibration is achieved. In fact, many of the figures referred to in the previous paragraphs also show tests which are well-behaved. Additionally, other series of tests not shown here displayed rapid equilibration. The discussion is intended to make the point that the equilibration process is sometimes not as straightforward as one would have expected, even when distilled water or standard aqueous leachate is used as permeant, and that in some cases (as in Figure 21) even initial behavior that is very stable may be followed by changes in permeability which are large relative to the initial variability exhibited by replicated specimens.

These writers have not been successful in determining which clays are most liable to exhibit irregular behavior in the equilibration process, based on the descriptors of the soils which have been provided. However, in private communications with Daniel, he has made the observation that the stucture of the soil after the moisturization step appears to affect ability to prepare uniform specimens which lead to well-behaved tests. According to Daniel, clays which tend to aggregate into large clods that are broken up only with great difficulty lead to inconsistent behavior. This would indicate that this stage may be deserving of special attention during specimen preparation.



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 30. Permeability of Hoytville Illite with .01N $CaSO_4$

Peirce's Faceville Clay series of tests, together with the other examples discussed in this Section, have a number of implications which are discussed below:

a. In comparative studies, it is advisable to include control tests in order to provide statistical confirmation of the effect of chemical permeants relative to water. Had such control tests not been provided in the series of Faceville Clay tests in Figure 21, it is reasonable to suppose that most observers would have concluded that the addition of chemicals had been the cause of the noticeable increase in permeabilities over the baseline levels established during water permeation. In fact, what appears to have occurred is that the effect is due to other factors. Without control tests, the possible effect of chemicals would have been confounded with the equilibration process, or conceivably with some unknown spurious factor which may have caused the observed increase.

   In addition to the Faceville Clay series, there are several examples of chemical test series in the data base where equilibrium was not clearly established and in which interpretation of the observed results was ambiguous, because control tests were not included in the experiments.

   There may be a tendency to dismiss these examples because the observed permeability increases were small relative to those caused by the introduction of chemicals in many such tests in the data base. Nevertheless, inclusion of control tests removes much of the possibility for ambiguity in attributing causality to the results of testing, even in cases where the observed effects are more pronounced. In any case, the increases observed in the Faceville Clay series would not seem to be inconsequential in terms of an application in which it was desired to determine whether a certain clay met current EPA standards of acceptability for liner material. A clay whose measured permeability was $10^{-7.6} = 2.5\times10^{-8}$ cm/sec might well be considered completely acceptable, whereas one whose permeability was $10^{-7.15} = 7.1\times10^{-8}$ cm/sec might be considered uncomfortably close to the maximally acceptable limit of 10^{-7} cm/sec (42). Alternatively, looked at from a design perspective, a .44 order of magnitude increase in permeability would require about a 66% increase in the liner thickness required to withstand breakthrough over a fixed time period.

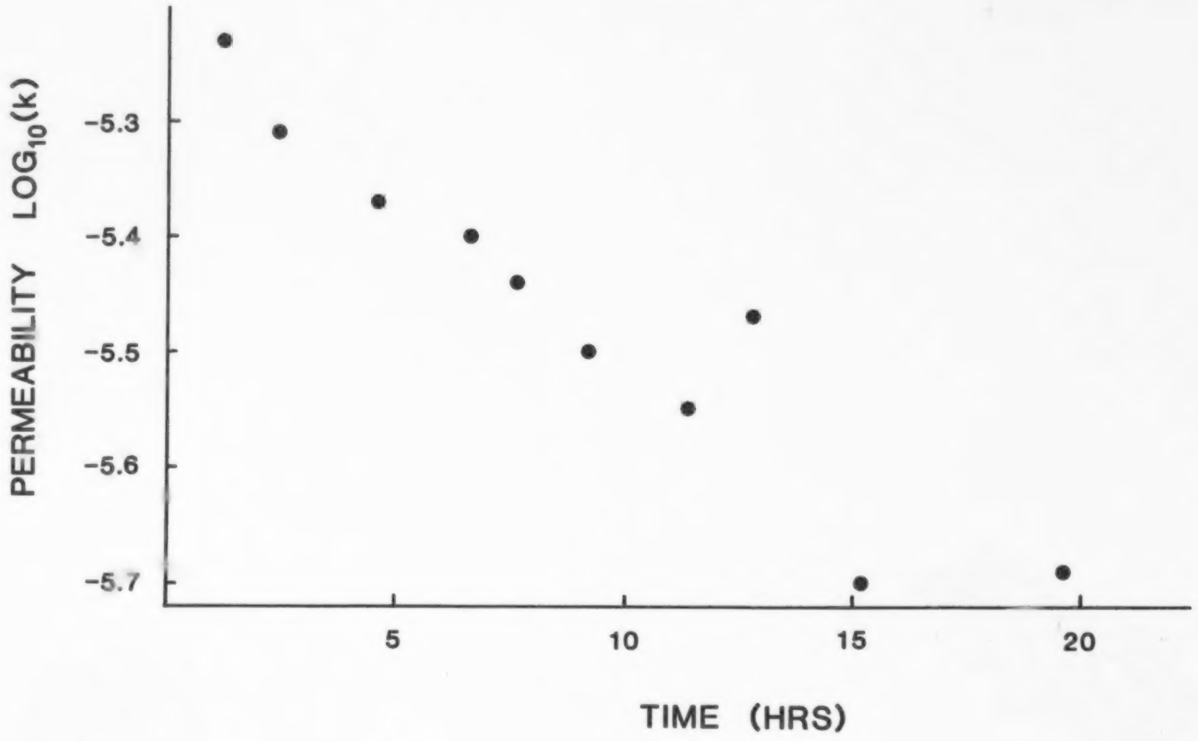
b. All permeability test results provided by private geotechnical firms used water as the permeant. They were almost always very much shorter than those made available by EPA contract researchers. Of 75 tests from private sources, none was longer than 146.2 hours. In terms of pore volumes the longest was 1.17 pore volumes, and only twelve exceeded .3 pore volume. Median length of testing was 53 hours (.16 pore volumes) for rigid-wall tests, and 47 hours (.04 pore volumes) for flexible-wall tests.

   In view of the long time required for many of the research tests to equilibrate, it is probable that tests as short as those described above will sometimes be considerably less precise than those provided from research sources. Unfortunately, there was almost no replication of tests on samples drawn from the same exact location and molded at identical water contents in the data provided by geotechnical firms. Thus variability due to the test itself cannot be differentiated from variability caused by location-to-location heterogeneity of soil samples, which may be considerable (see Section 4.2). Although a reasonably large number of private sector tests had clearly not equilibrated prior to termination, in other cases their behavior did appear to be stabilized (in many cases the number of repeated determinations was only two or three, making any assessment of equilibration impossible). Even so, the examples reviewed in this Section indicate that the initial behavior of water tests over relatively short time intervals may not accurately predict their final behavior, as illustrated by Figure 21. While Figures 21-30 show many tests whose terminal permeabilities differ only negligibly from those obtained over the first 100 hours of testing, they also show many whose initial permeabilities differ by half an order of magnitude or more from the levels obtained in the long run.

The most common pattern followed by the flexible-wall tests from the private sector was one of decreasing permeability over time (although in two cases permeability increased by as much as one order of magnitude during the test period). Figure 31 shows one such test whose equilibrium at the end of the test period has not been established. Its permeability fell about a half-order of magnitude over 20 hours of testing. Decreasing permeability over time was also the norm for those flexible-wall water tests from EPA contract researchers which experienced any appreciable changes from their initial levels (see Figures 24 and 30). This point may be of practical importance in the evaluation of proposed liner material, because it suggests that short-term water tests using flexible-wall permeameters would usually yield results which conservatively overestimate their terminal levels. This assumption seems reasonable since it would be expected that the confining pressure applied to a flexible-wall sample would initially consolidate the soil to some degree.

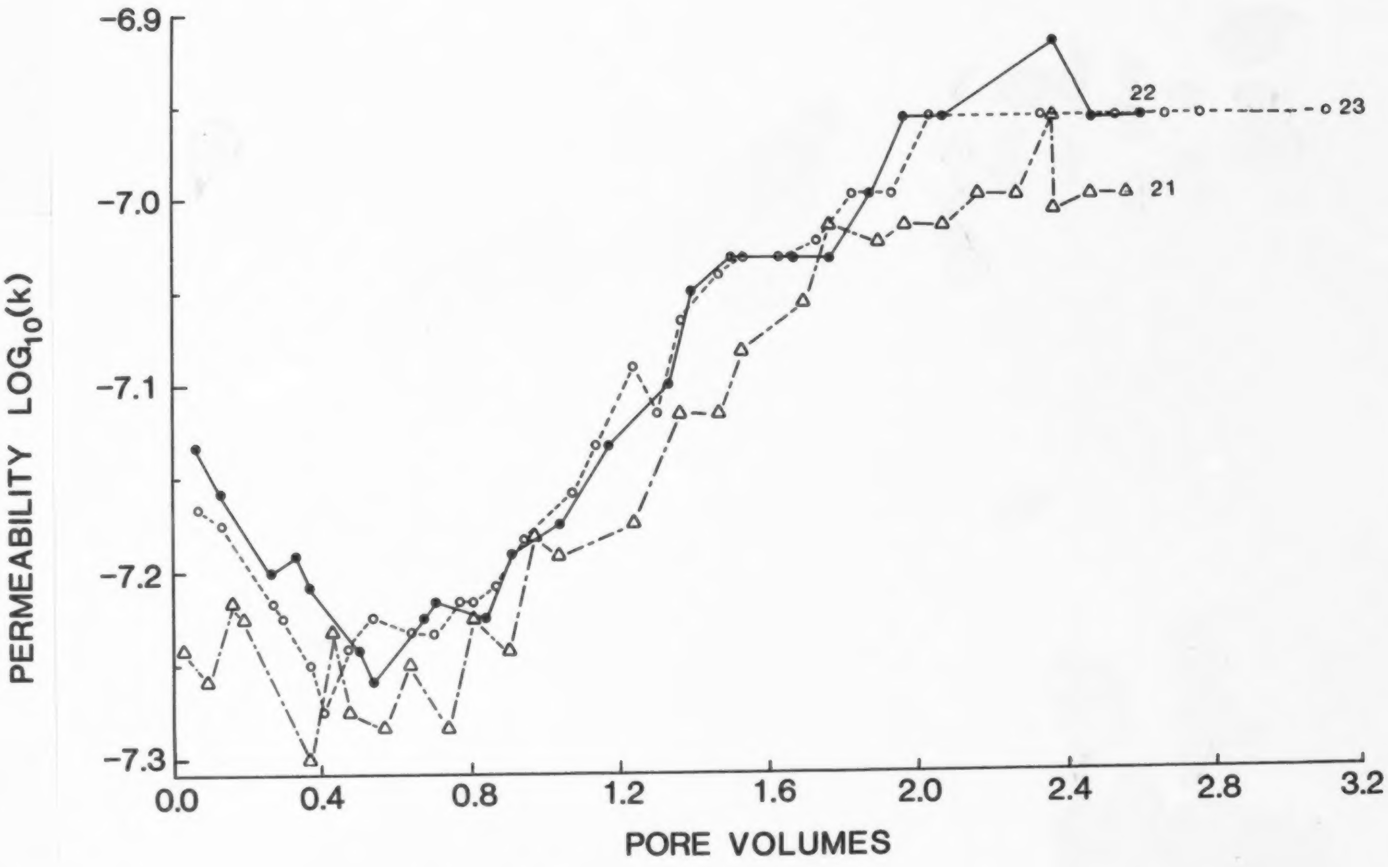
The phenomenon of decreasing permeability over time with flexible-wall permeameters was not seen to be typical in the case of tests employing chemical leachates. See for example Figure 32, which shows replicated flexible-wall tests of methanol on Kaolinite Clay. These tests are due to Daniel.

In the flexible-wall tests obtained from the private sector which used water as the permeant, it was a common practice to stop tests overnight (these tests used the falling-head method). Figure 33 shows one such test. A noticeable pattern is evident in which permeability seems to be decreasing in steps, with the total decrease during the test period being nearly one order of magnitude. The major drops in permeability appear to be occurring overnight, during which time the flow of water was stopped. The first four observations were made on a Wednesday, the second four on Thursday, and the



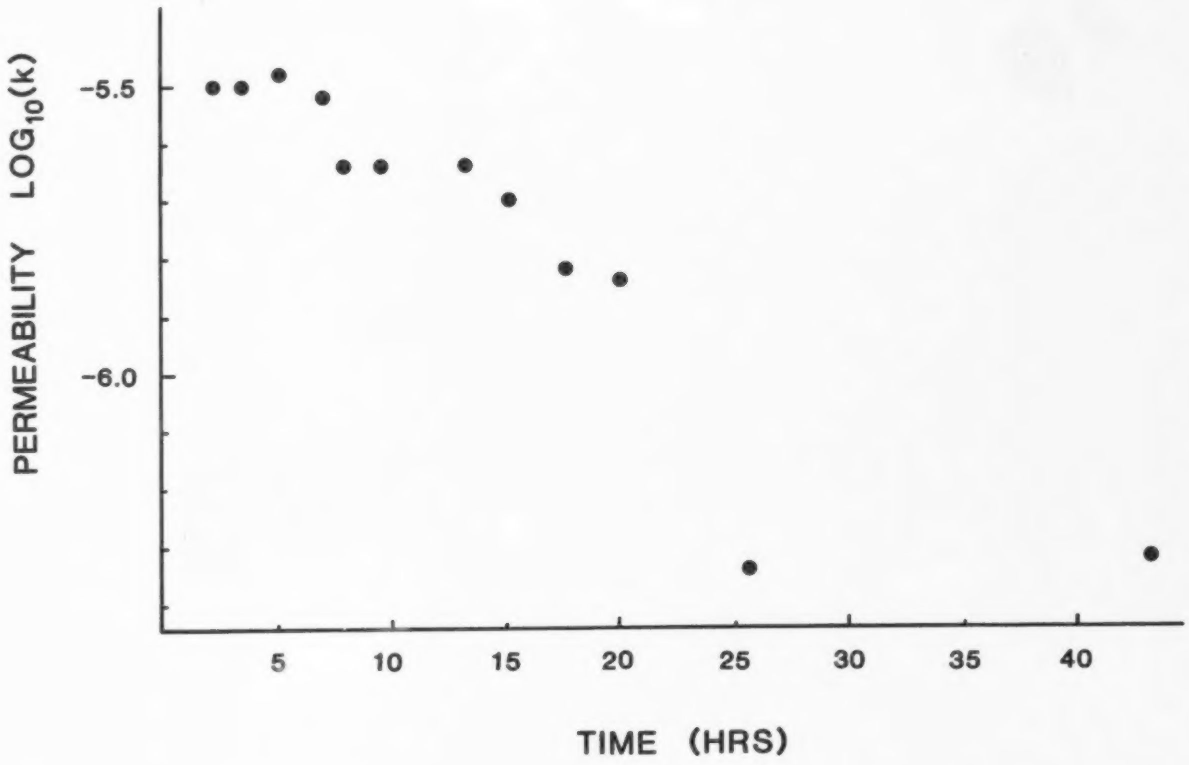
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 31. Decrease in permeability of a flexible-wall sample with water



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 32. Permability of Kaolinite Clay with methanol (flexible-wall test with i = 50)



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 33. Apparent effect of consolidation on a flexible-wall sample with water

following two on Friday. The last two observations were made on Monday and Tuesday of the following week. The test was stopped during each overnight period, except for the one separating the last two observations. Thus it appears that noticeable consolidation is occurring during those periods in which the test has been suspended. It is conjectured that in this test the system may have been depressurized during the overnight periods, which could considerably increase the consolidation stress applied by the sidewall pressure. This would lead to a more densely consolidated specimen and therefore to a spuriously low terminal permeability.

4.5 EFFECT OF GRADIENT

Daniel and Brown have provided tests of clay soils which were conducted at several different gradients, thereby allowing an investigation of whether this factor has a significant effect on the permeability obtained by use of Darcy's Law.

In one series of experiments, Daniel prepared 24 specimens of Kaolinite Clay at nominally equal molding water contents of about 33%, or about 2% above optimum water content. From Figure 13 this appears to be very near the point at which the permeability-molding water content relationship is minimized, so that the effect of molding water content on permeability should be minimal.

Twelve of the specimens were permeated with water (.01N $CaSO_4$), three at each of the gradient settings 10, 50, 100 and 300. The remaining twelve specimens were similarly permeated with 100% methanol. Flexible-wall permeameters were used in the experiments, with an effective confining stress of 15 psi applied to all tests except those at a gradient of 300, where an effective stress of 25 psi was required to keep the membrane pressed against the soil.

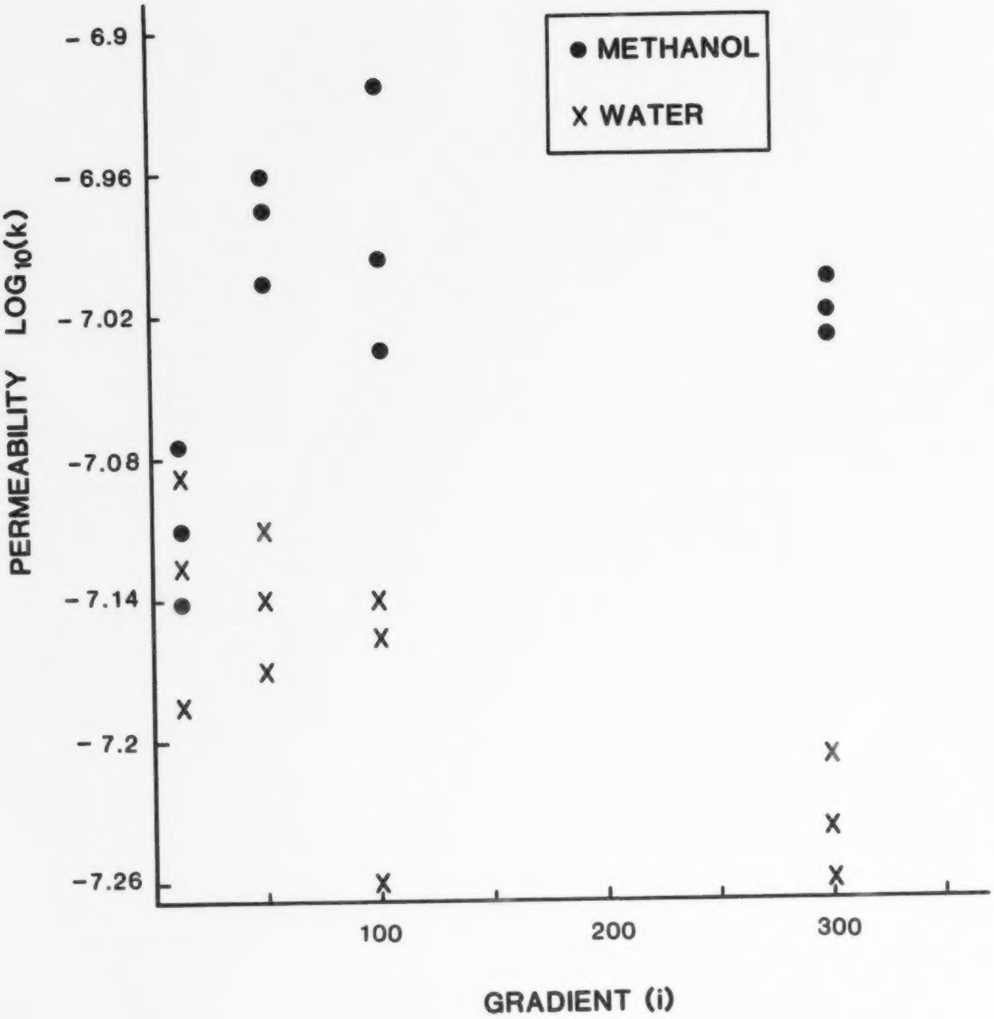
Tests were run for at least 2 pore volumes and in some cases more than four pore volumes. The tests which employed methanol typically increased in permeability by about .3 of an order of magnitude over the course of the experiment, but by the termination of the experiment steady behavior of all tests appears to have been achieved (Figure 32 shows those tests on methanol at a gradient of 50, and is typical in appearance to the others).

Figure 34 shows the terminal permeabilities of each of the twenty-four tests, which were obtained by applying Darcy's Law to the flow observed in each test over the last half-pore volume (or slightly more, depending on when individual recordings were made). In Figure 34 results of tests on methanol are denoted by black circles and those on water by "X's".

Several features of Figure 34 are noteworthy:

1. The final permeabilities with methanol are about .2 of an order of magnitude (58%) above those with water, except for those tests run at a gradient of 10, where there seems to be little difference.

2. Except for the methanol tests at a gradient of 10, permeability appears to be perhaps slightly decreasing as gradient increases,



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 34. Flexible-wall tests on Kaolinite Clay with .01N $CaSO_4$ and methanol at gradients of 10, 50, 100, and 300

although the magnitude of the effect is very slight.

3. The within treatment (chemical -gradient combinations) variance of log permeabilities is extremely small. The standard deviation is only .04, with no apparent difference in variability between the water tests and the methanol tests. This data set was one of the least variable that was investigated.

4. Molding water content is not significantly related to the within-treatment variation of log permeabilities, as tested by analysis of covariance methods. This is consistent with statements made in the second paragraph of this Section.

5. If formally tested by analysis of variance (or analysis of covariance) methods, the difference between methanol and water is highly significant ($p=.0001$) and the effect of gradient within chemicals is highly significant ($p<.01$) with strong interaction between the two factors ($p=.003$) caused by the surprising location of the permeabilities to methanol at a gradient of 10. This is because, although the magnitude of the observed effects is rather small from a practical point of view, it is not small relative to the variation of log permeabilities within treatments.

These writers believe that the results of the methanol tests at a gradient of 10 are very likely due to causes other than gradient effect and that the formal statistical tests cited above give misleading results because the effect of gradient has been confounded by other factors whose effects may be relatively small, but large enough in this situation to lead to spurious results.

According to the experimenter (private communications with Daniel), tests were prepared in the order that results were filed in the data set. Therefore, potential biases were not randomized out of the experiment, because the order of listing started with the methanol tests at a gradient of 10, and continued in this way until the final listing of the water tests at a gradient of 300 was completed. For this reason, a number of factors could be confounded with both gradient and chemical comparisons, including variation of soil characteristics and improvement or other differences in preparation and compaction technique. This confounding would be exacerbated by the fact that multiple specimens were sometimes compacted from commonly moisturized batches of the clay soil. If the tests within individual chemical-gradient combinations were prepared for the most part from the same moisturized batch, then the within-treatment variance of log permeabilities reflects primarily within-batch variation, whereas comparisons of treatments generally would involve an additional component of between-batch error. The hypothesis that this may be occurring in this data set is strengthened by the fact that the within-treatment variance of log permeabilities is so small.

Because the range of gradients was so large in this experiment, the time over which experiments were run varied systematically. That is, the tests at a gradient of 10 would have been run somewhere in the neighborhood of 2700-3800 hours to accumulate the number of pore volumes reported, whereas the tests at a gradient of 300 should have required roughly 200 hours (the ratio of times is not too close to 300:10 because the tests at higher gradient were allowed to accumulate more pore volumes). This could permit biasing effects if environmental conditions within the lab changed by meaningful amounts over the course of the experiment or if the pressure applied to concurrently run tests varied slightly from its nominal value in a way that might induce correlation among them.

The belief that the pattern exhibited in Figure 34 may not be attributable to gradient differences may be partially checked. Daniel ran several additional tests using water or methanol as permeant on Kaolinite Clay soil at molding water contents within the range spanned by those of the twenty-four tests in Figure 34. These tests were all much shorter in terms of pore volumes of flow than the tests of Figure 34, but at least two of them appear to have achieved equilibrium.

Shown in Figure 35 is a time series plot of a test using methanol on Kaolinite Clay at a gradient of 2. After about 150 days of testing it has accumulated .64 pore volumes and appears to be stable over the last 1500 hours of the test (.29 pore volumes). Its average permeability during that time is 1.04 x 10^{-7}cm/sec, or -6.98 on the log scale.

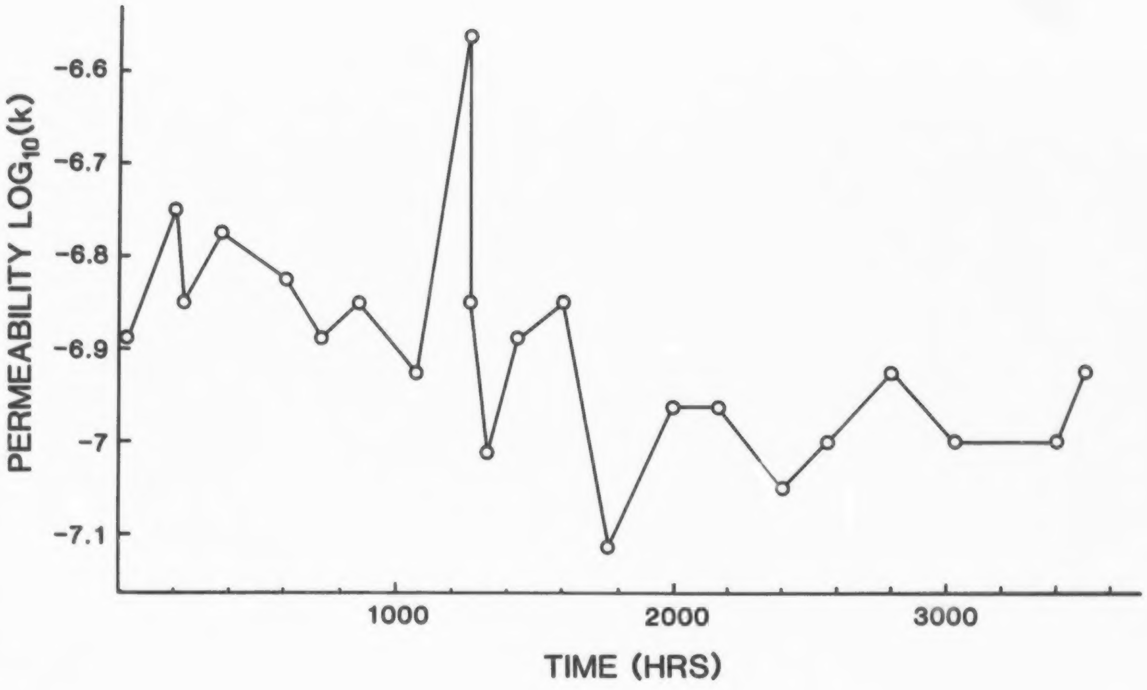
In Figure 36 is a time series plot of a test run at a gradient of 30, using .01N $CaSO_4$ as permeant. It has run about 40 days, or 1.28 pore volumes. Its permeability appears to have stabilized over the last two thirds of the test and averages 5.98 x 10^{-7} cm/sec = $10^{-7.26}$ over the last half pore volume.

The above two additional permeability values are added in Figure 37. They agree quite well with the previously discussed tests in terms of the amount of variability usually found in permeability testing (e.g. all water tests are in agreement to within .2 of an order of magnitude, and the methanol tests agree to within .22 orders of magnitude). However, their placement provides an indication that the earlier tests were influenced by within-treatment correlation. Most noticeably, all three tests on methanol at a gradient of 10 seem out of place and possibly have been influenced by some common factor.

To these writers, this data set indicates that for this clay soil and the permeants investigated, the effect of gradient is not well established and, in any case, is quite moderate over the range investigated, with perhaps a very slight decrease in average permeability as gradient increases. This could reasonably be attributed to increased consolidation due both to higher gradients and (for i=300) to the increased effective confining stress.

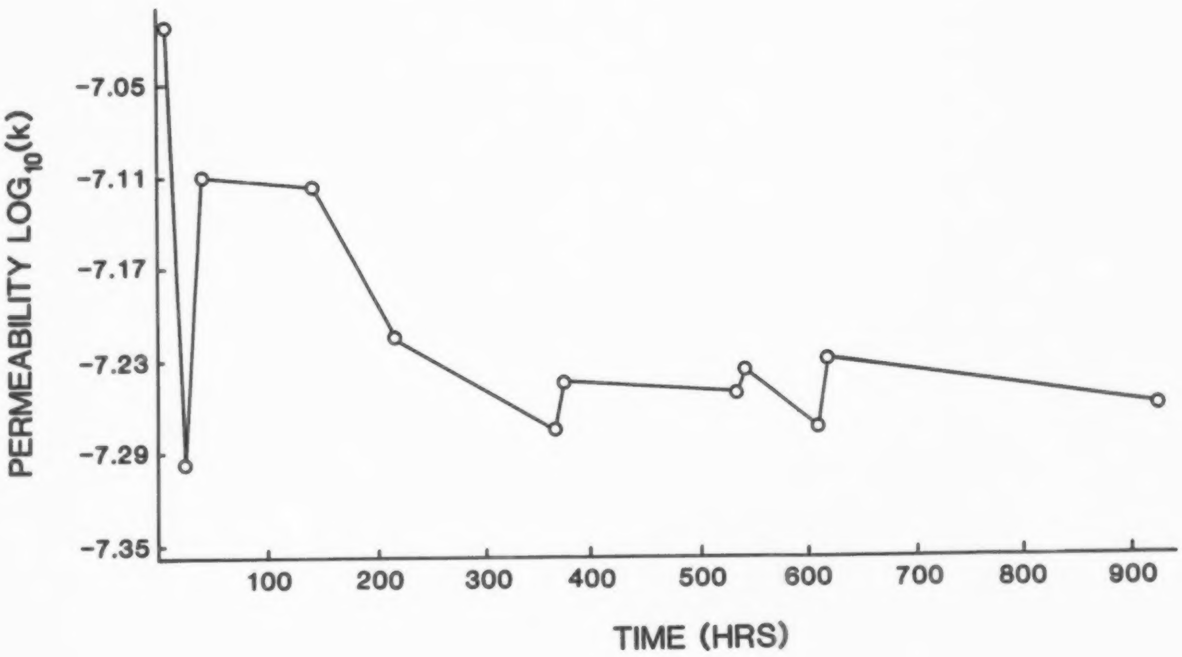
Other tests conducted by Daniel indicate that in some circumstances, the gradient at which a test is run may have a much more pronounced influence on its nominal permeability than was the case in the Kaolinite experiments. Tests were also prepared using Lufkin Noncalcareous Smectite and Hoytville Illite, using both methanol and .01N $CaSO_4$ as the permeant liquid. These tests were run in flexible-wall permeameters at gradients of 10, 50, 100, and 300.

Those tests in which .01N $CaSO_4$ was permeated through Lufkin Noncalcareous Smectite are not considered here, because in some cases they did not appear to have equilibrated completely. The general pattern of these tests was one of decreasing permeability over time, and it was thought that their



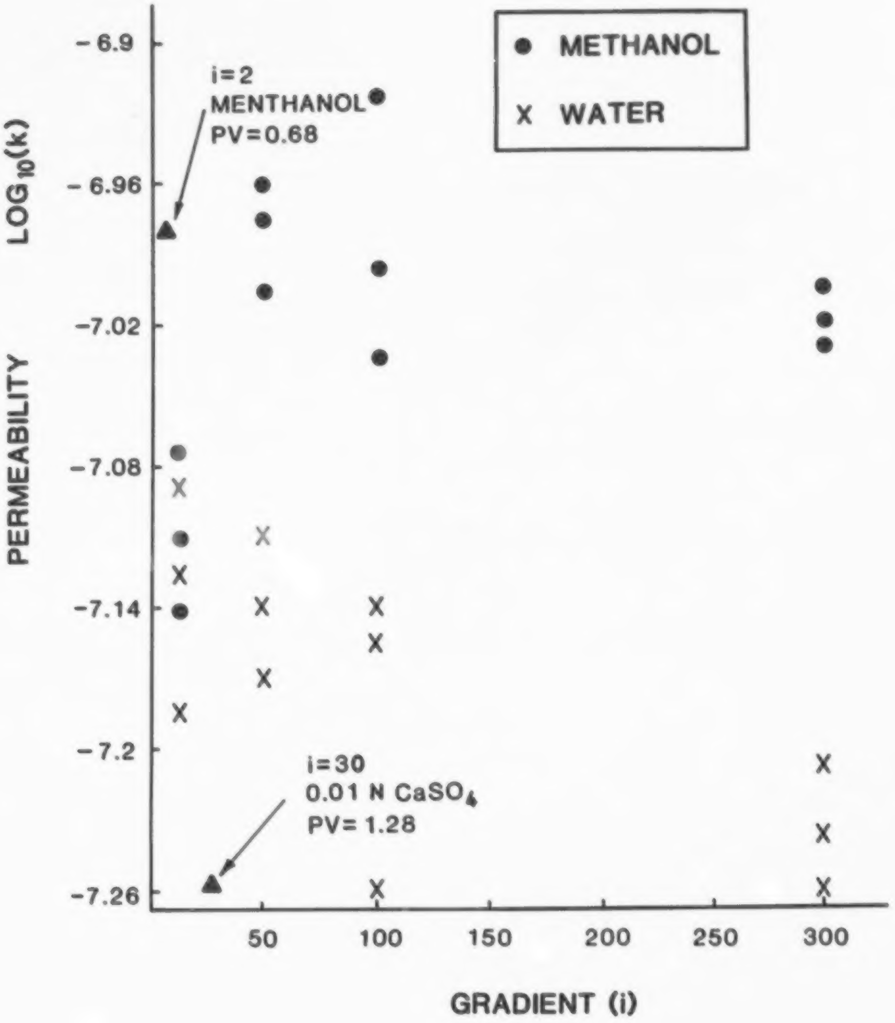
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 35. Permeability of Kaolinite Clay with methanol (flexible-wall test with i = 2)



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 36. Permeability of Kaolinite Clay with .01N $CaSO_4$ (flexible-wall test with i = 30)



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 37. Flexible-wall tests on Kaolinite Clay with .01N $CaSO_4$ and methanol with gradients of 2, 10, 30, 50, 100, and 300

inclusion might lead to the appearance of a gradient effect which could be explained by the fact that in terms of real time, the tests at higher gradients might equilibrate more quickly than those at lower gradients. This is illustrated in Figure 38, which shows permeabilities of several of the shorter tests on Lufkin Noncalcareous Smectite with .01N $CaSO_4$ graphed against what might be termed "adjusted" time, equal to real time multipled by the gradient at which a test is run divided by 100. This "adjusted" time attempts to compensate for the acceleration or deceleration of the tests due to gradient differences, and uses a gradient of 100 as baseline.

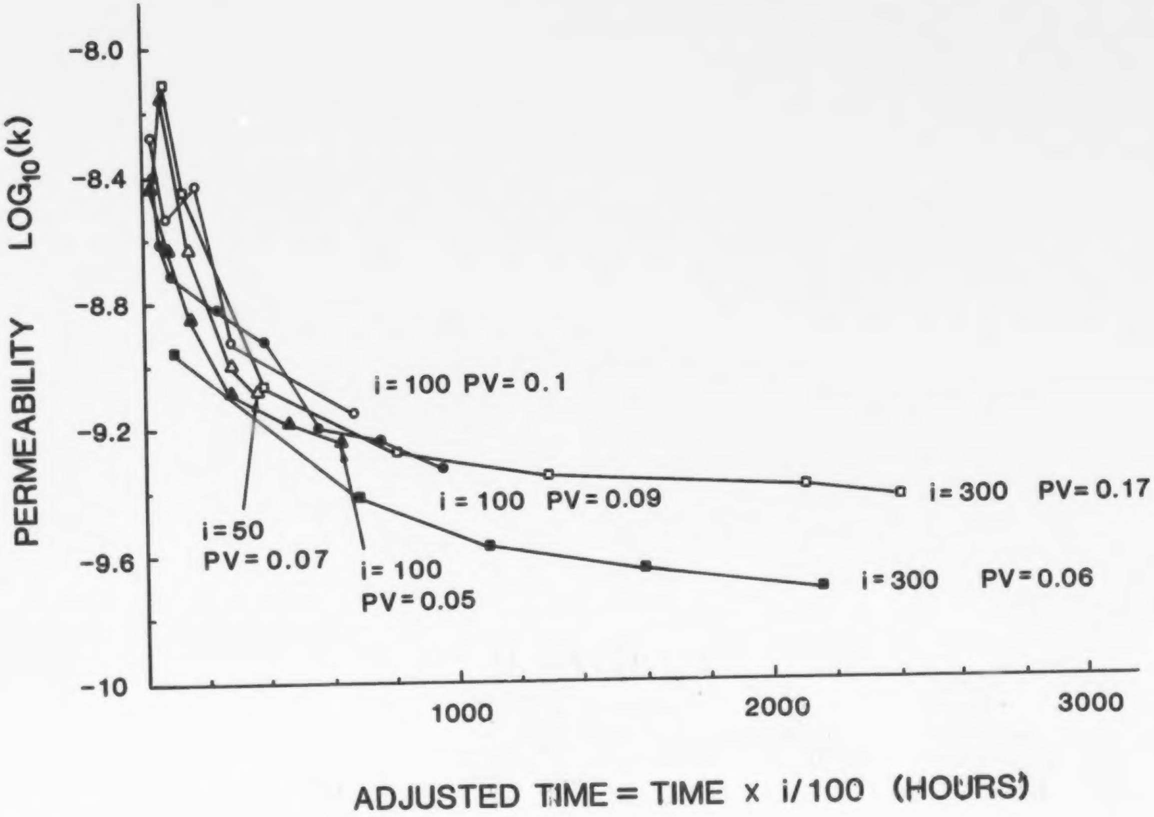
In Figure 38, note that the tests have similar trends, but that the tests which were run at a gradient of 300 have lower terminal permeabilities than tests at a gradient of either 50 or 100. In terms of real time, these tests were of about equal length, but the ones at a gradient of 300 are longest in terms of adjusted time and appear to have progressed towards equilibrium more than the others.

The tests on Lufkin Noncalcareous Smectite with methanol and the tests on Hoytville Illite with either permeant were generally much longer in terms of pore volumes than the tests in Figure 38, and do not appear to present a problem of biased gradient effects due to differences of effective time on test. See Figure 39 as an illustration of this.

Figures 40, 41 and 42 show the terminal permeabilities obtained on Lufkin Noncalcareous Smectite with methanol, Hoytville Illite with methanol, and Hoytville Illite with water, respectively. Permeabilities in these figures are based on the flow within each test over the last half-pore volume of the experiment, with the exception of the two tests on Hoytville Illite with water at a gradient of 10. These tests were short in terms of pore volumes (.19 and .55 pore volumes) and so their terminal permeabilities were based only on the last .1 pore volume (in each case in excess of 1700 hours), over which they appeared to be in equilibrium.

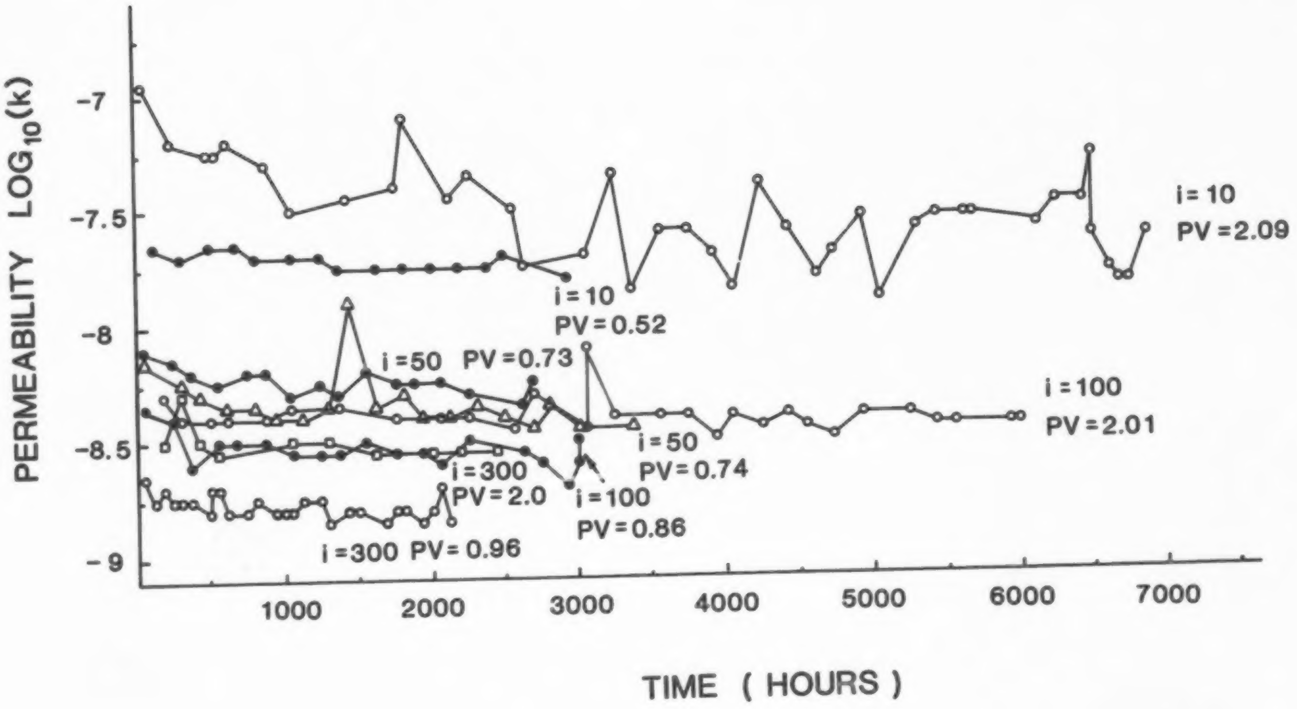
Figure 40 indicates a strong gradient effect for the Lufkin Noncalcareous Smectite with methanol, with the specimens tested at a gradient of 10 having permeabilities one-half to one order of magnitude greater that those tested at higher gradients. In Figure 41, the tests of Hoytville Illite with methanol show a similar pattern, although not as convincingly. In Figure 42, the Hoytville Illite with water seems to have only slightly decreased permeability as gradient increases.

Tests on the Hoytville Illite were run at molding water contents ranging from 17.8 -19.2% (Optimum Moisture Content = 18%), and variation in molding water content appears to account for some of the observed differences in the permeabilities of the Hoytville Illite specimens. In Figure 43, the permeabilities are re-plotted after adjusting the Hoytville Illite tests for differences in molding water content by analysis of covariance. The Lufkin Noncalcareous Smectite tests are not adjusted in this plot since, after discussion with Daniel, it was decided that three of the reported water contents (one at each of the gradients 10, 100 and 300) might not be reliable, as they did not correlate well with the observed densities or permeabilities. However, these tests have permeabilities which are in good agreement with their replicates, and in any case covariance adjustment of the Lufkin



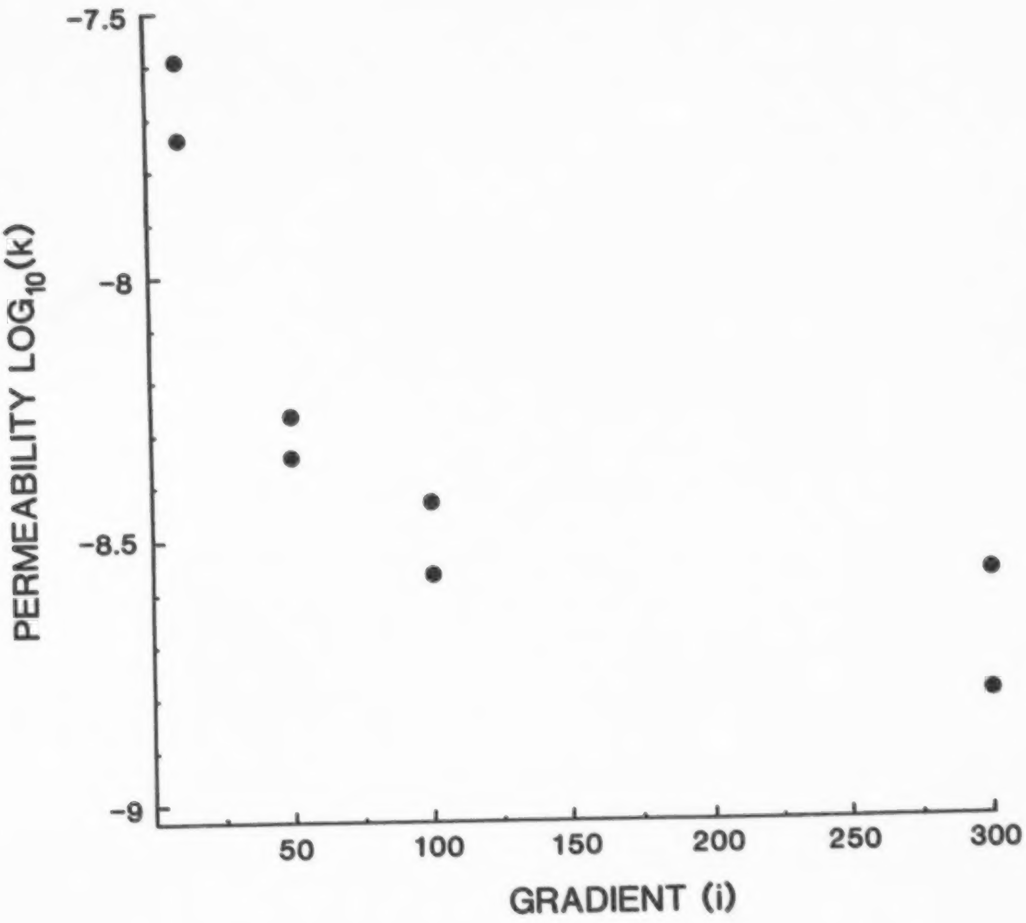
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 38. Permeability of Noncalcareous Smectite with .01N $CaSO_4$ (flexible-wall tests)



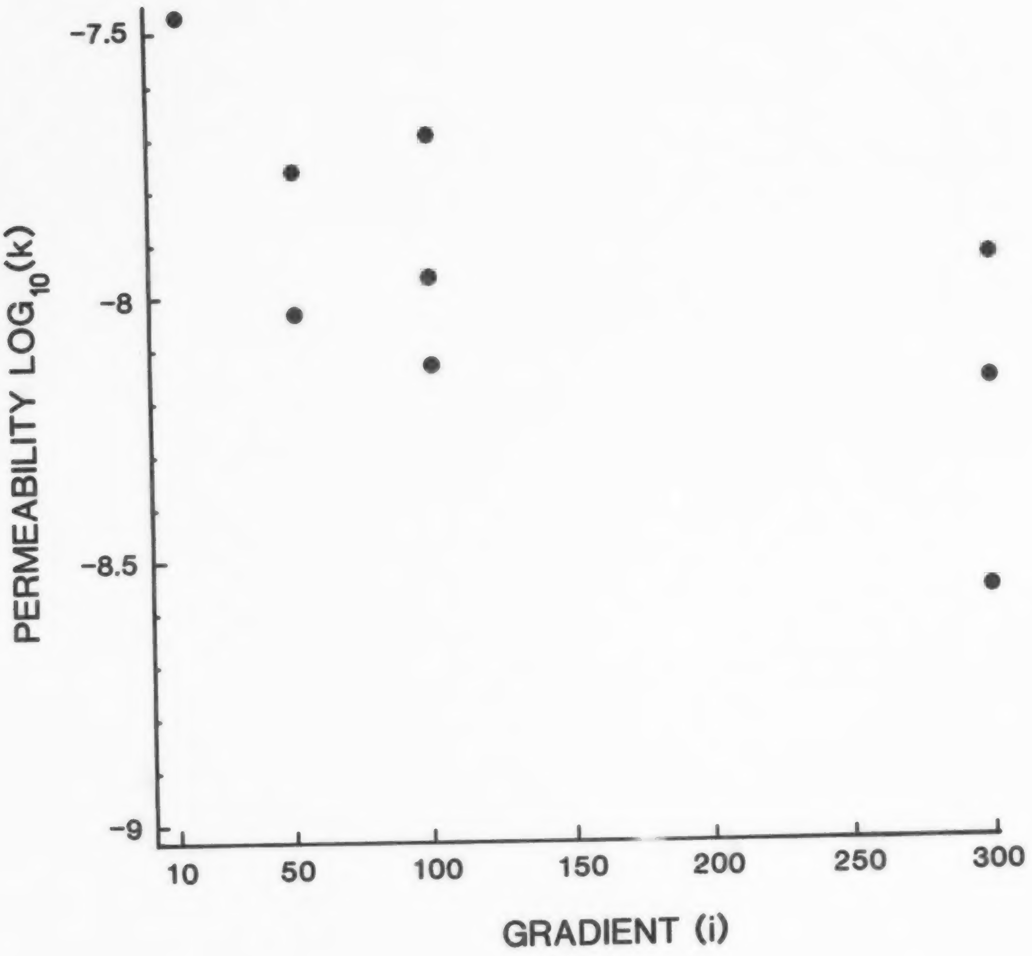
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 39. Permeability of Lufkin Noncalcareous Smectite with methanol (flexible-wall tests)



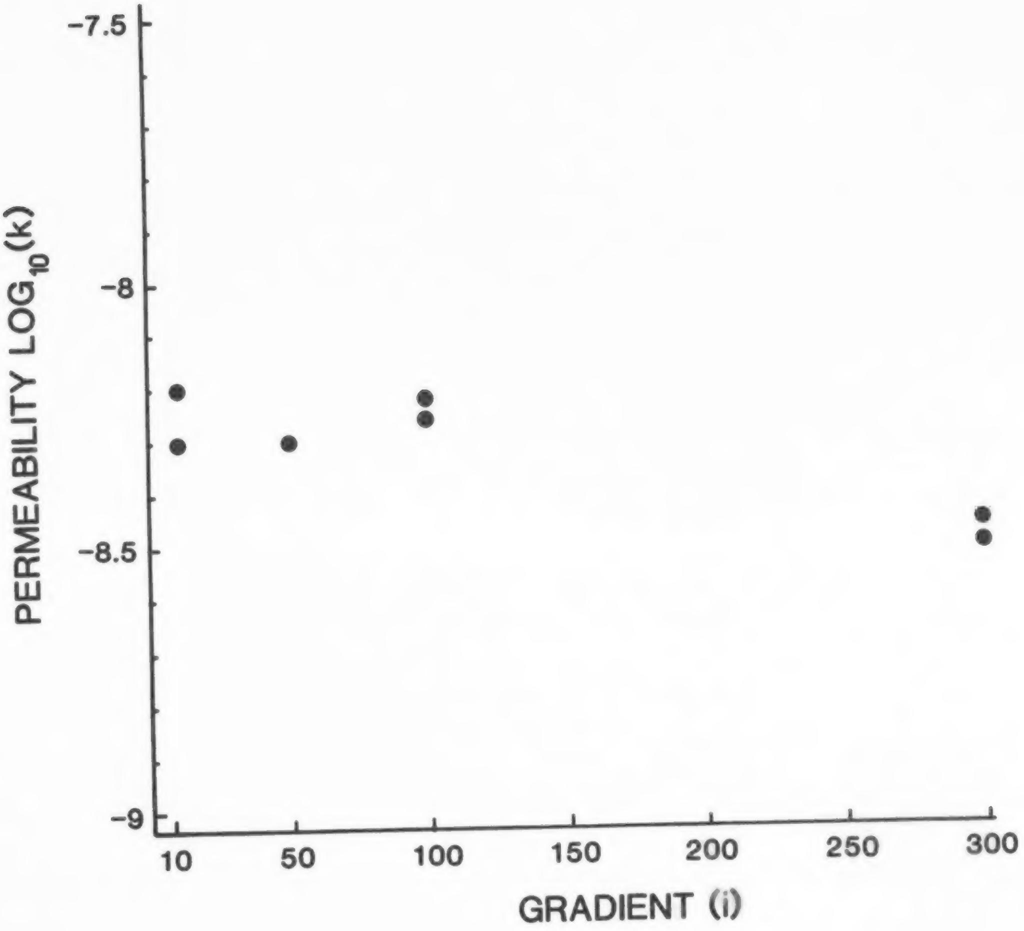
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 40. Effect of gradient on permeability of Lufkin Noncalcareous Smectite with methanol (flexible-wall tests)



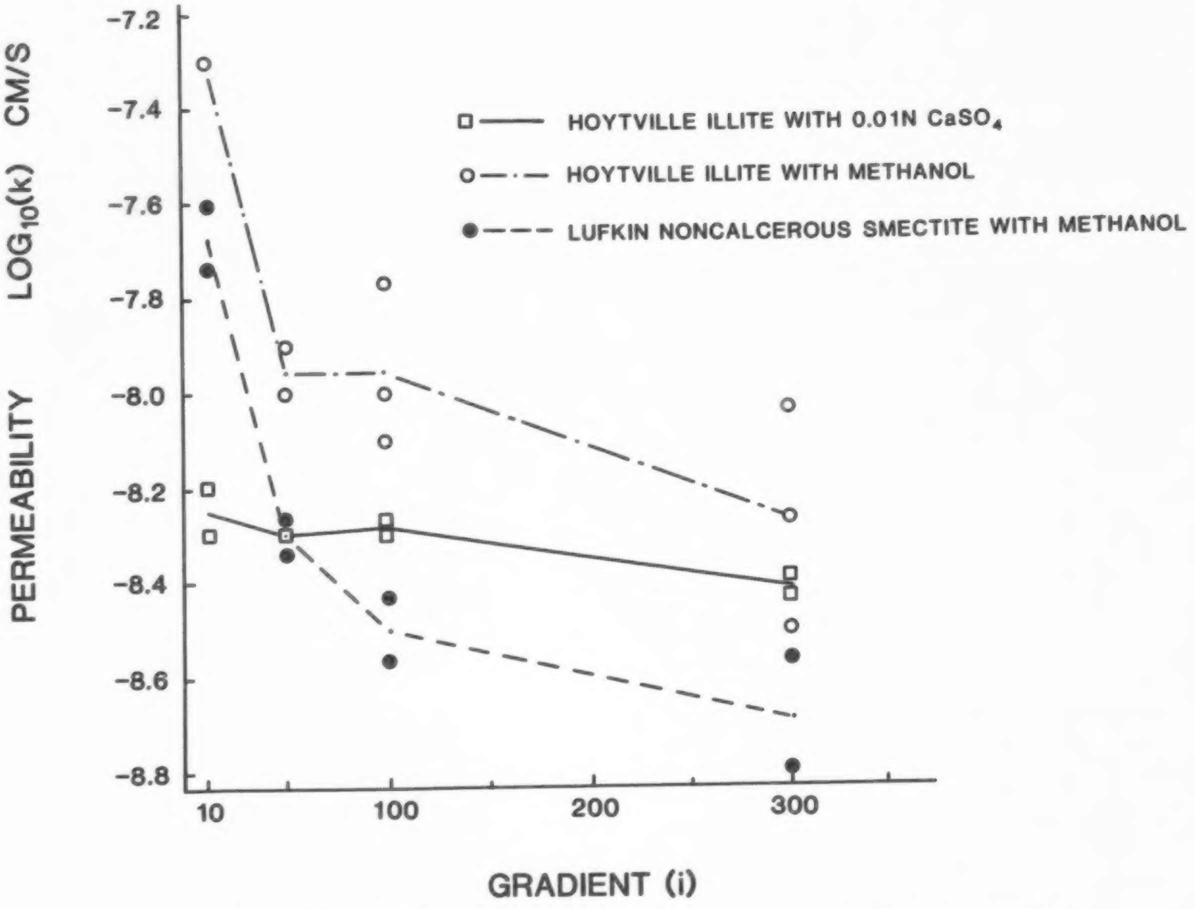
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 41. Hoytville Illite with methanol (flexible-wall tests)



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 42. Hoytville Illite with .01N $CaSO_4$ (flexible-wall tests)



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 43. Effect of gradient on the permeability of two clays (Hoytville Illite permeabilities adjusted for differences in molding water content as discussed in text)

Noncalcareous Smectite permeabilities using the reported data leads to virtually no change in the appearance of Figure 43.

In Figure 43, the strong effect of gradient appears similar in the test series on both clays which use methanol as permeant. The apparent effect of gradient in the Hoytville Illite specimens which have been permeated with water is very minor. This impression is reinforced by formal statistical tests, which strongly indicate a main effect due to gradient ($p=.0002$) with a more moderate indication of gradient-series interaction ($p=.04$) caused by the dissimilarity of the Hoytville Illite-water profile in Figure 43 relative to the other two. Within series, the effect of gradient in the Lufkin Noncalcareous Smectite-methanol series is strongly indicated ($p=.0002$) as is also the case in the Hoytville Illite-methanol series ($p=.003$). No gradient effect is apparent in the Hoytville Illite-water series ($p=.91$).

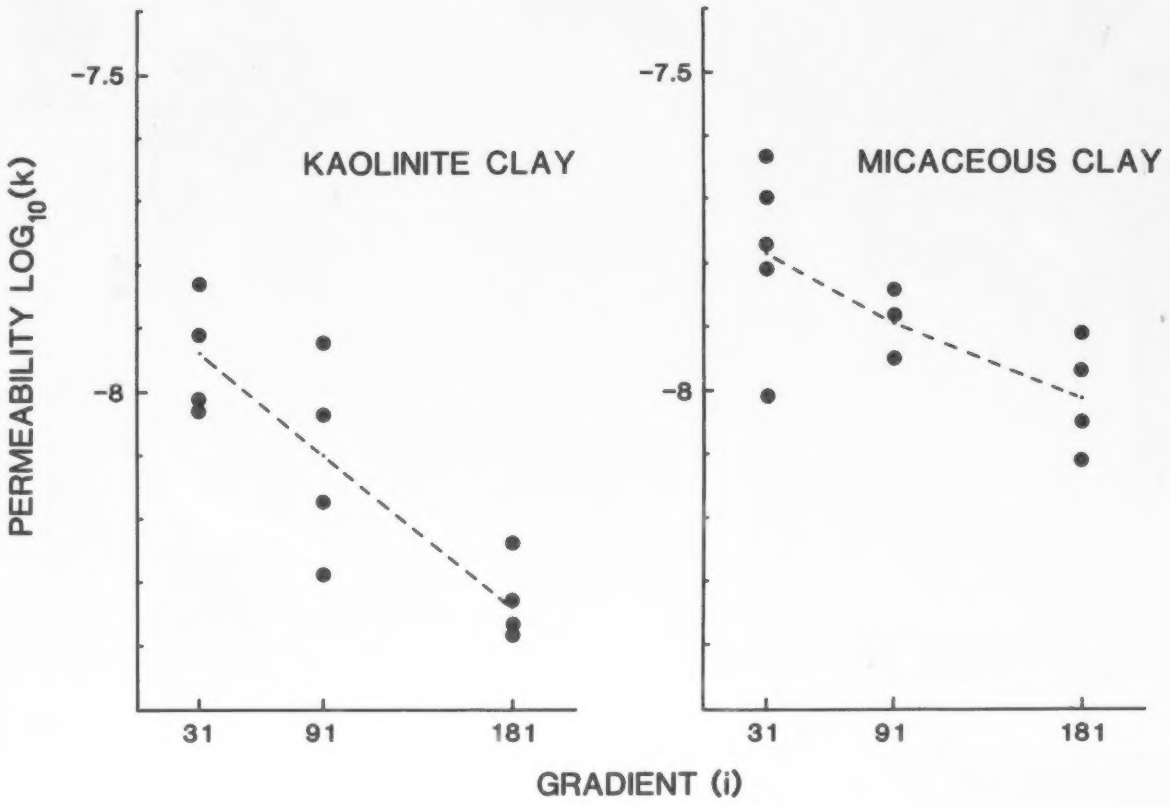
It does not appear reasonable to attempt to attribute these effects to spurious causes, as was suggested in the discussion of Figure 34. This is primarily due to the greatly increased magnitude of the observed effects. Furthermore, the tests in Figure 43 were generally not compacted nor run in a systematic order, thus reducing the likelihood that batching effects may have been confounded with gradient.

Brown has provided rigid-wall tests of Elgin Kaolinite Clay, Micaceous Clay, and a Bentonite Clay using .01N $CaSO_4$ as a permeant. These tests were conducted at gradients of 31, 91 and 181 for the first two soils, and at gradients of 91, 181 and 272 for the Bentonite Clay. In most cases the tests passed about one pore volume before termination, but in some cases 1.5 -2.0 pore volumes were passed, and in one case only .6 pore volume was passed.

Shown in Figure 44 are the permeabilities of the Elgin Kaolinite Clay and Micaceous Clay tests. These were computed based on the total flow over the final .4 pore volume of the tests, since several of them appeared to be trending over time somewhat prior to that point (however, basing the calculation on the last .2, .3 or .5 pore volumes turned out to have little effect on the computed permeabilities). One of the Elgin Kaolinite Clay tests provided was not included in this analysis, since its behavior was quite erratic (see Figure 28).

In Figure 44, permeabilities of both soils decrease as gradient increases. On average, the permeability of the Elgin Kaolinite specimens at a gradient of 31 was .38 orders of magnitude (140%) larger than it was at a gradient of 181. The corresponding increase for the Micaceous Clay specimens was .22 orders of magnitude (66%), although analysis of variance gives no indication of interaction between gradient and clay type (i.e., supports the hypothesis that the effect of gradient is the same for each clay soil). The main effect of gradient is significant at the .0001 level.

In most cases, the molding water content at which these specimens were prepared was not available. However, the given data show little variation in the densities of the specimens, indicating that water content was probably controlled adequately. Furthermore the within treatment variation of log permeabilities was not significantly related to void ratio of the specimens.



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 44. Effect of gradient on the permeabilities of Kaolinite Clay and Micaceous Clay with .01N $CaSO_4$ (rigid-wall tests)

An attempted analysis of the Bentonite Clay permeability tests gave ambiguious results, with no clear pattern emerging. Several of these tests experienced significant decreases in permeability over time, presumably caused by swelling of the soil. In some cases this was followed by significant increase in permeability. It was not possible to assign a terminal permeability to several of the tests with enough confidence to justify any interpretation.

The results of the Elgin Kaolinite Clay and Micaceous Clay tests were somewhat surprising. In rigid-wall permeameters, these writers had expected to find, if anything, an increase in permeability due to gradient caused by the promotion of channel formation at higher pressures. In some chemical permeant tests in rigid-wall permeameters such an increasing effect may appear to be operative. Further discussion of this point will be deferred to Section 4.7. However, the water tests discussed in this Section seem to indicate that in some cases the dominant mechanism is one of increased consolidation at higher gradients, leading to decreased permeabilities.

Several of the test series discussed in this Section suggest that, in some circumstances, the gradient at which a test is run may have an effect whose magnitude is of practical significance and whose orientation is such that recorded permeabilities are negatively related to the gradient at which the test is run. It is not clear whether this result may be extrapolated to the lower gradients that would normally be expected in field conditions, but if such extrapolation is warranted then this would indicate that laboratory determinations run at higher gradients could understate the permeability to be expected in the field by sizeable amounts.

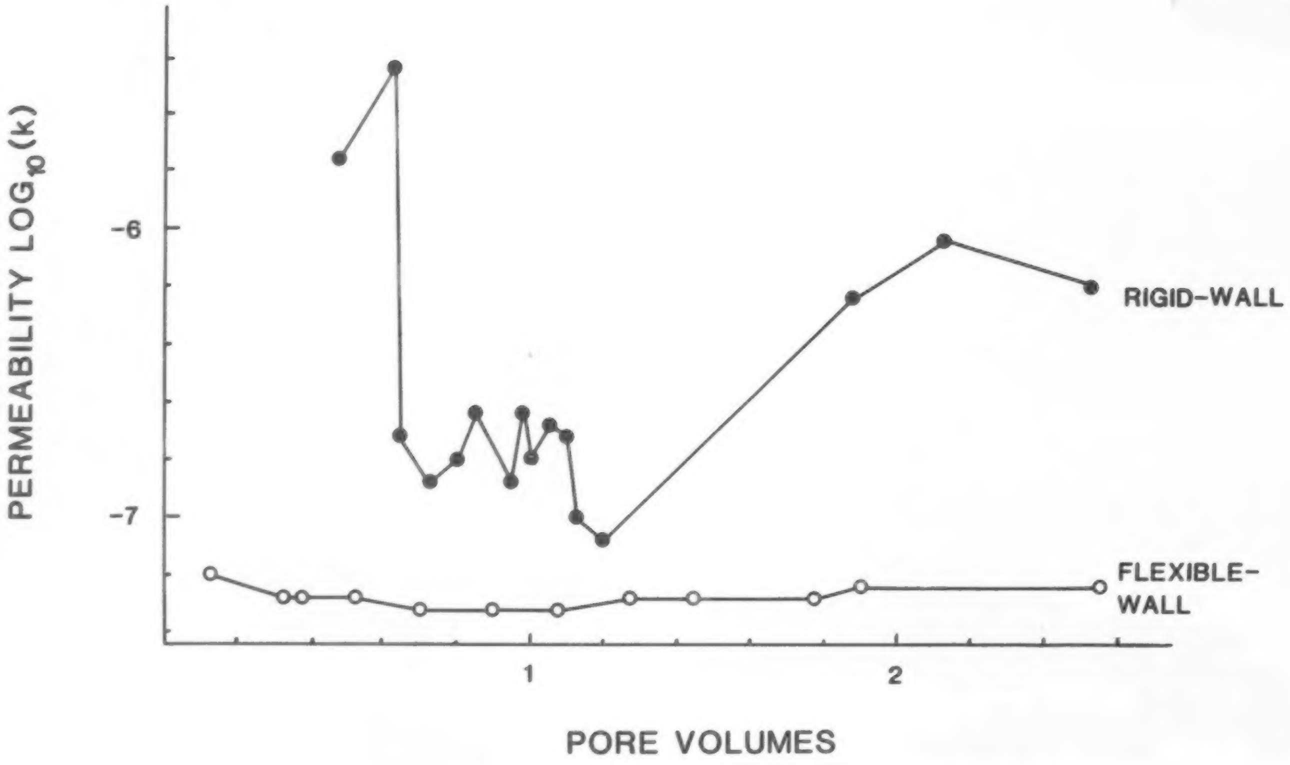
4.6 COMPARISON OF FLEXIBLE-WALL AND RIGID-WALL PERMEAMETERS

Tests on specific clay soil chemical permeant combinations, using both flexible-wall and rigid-wall permeameters, were included in the data provided by both Peirce and Daniel. In addition, several tests conducted by Brown and Anderson which used rigid-wall permeameters, made use of the same soils and chemicals that Daniel investigated in flexible-wall testing.

Matched within-lab data on the comparative effects of permeameter type on 100% concentrated chemical permeants was unreplicated. However, a review of the behavior of these tests is informative.

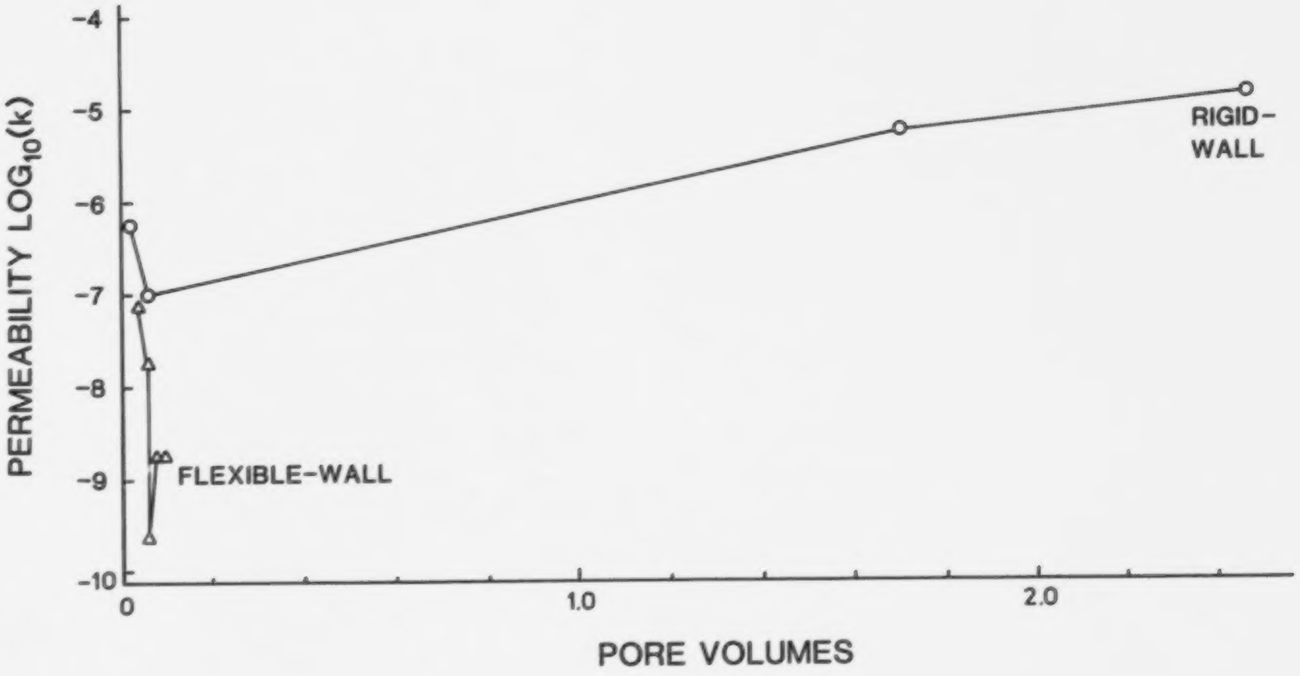
The results of comparable flexible-wall and rigid-wall tests on Kaolinite Clay using 100% acetic acid as permeant are shown in Figure 45. These are by Daniel, and were both conducted at a gradient of 200 and a nominally equal molding water content of about 33%. An effective confining stress 15 psi was applied to the flexible-wall sample. At termination the flexible-wall test permeability values stabilized at of about $10^{-7.28} = 5 \times 10^{-8}$ cm/sec over the last half pore volume while the rigid-wall test values were about one order of magnitude higher and may not have reached equilibrium.

Shown in Figure 46 are tests by Daniel of Kaolinite Clay using trichloroethylene in flexible-wall and rigid-wall permeameters. The behavior of the tests is fundamentally different; the rigid-wall test reached a



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 45. Kaolinite Clay with acetic acid in rigid-wall and triaxial flexible-wall permeameters



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

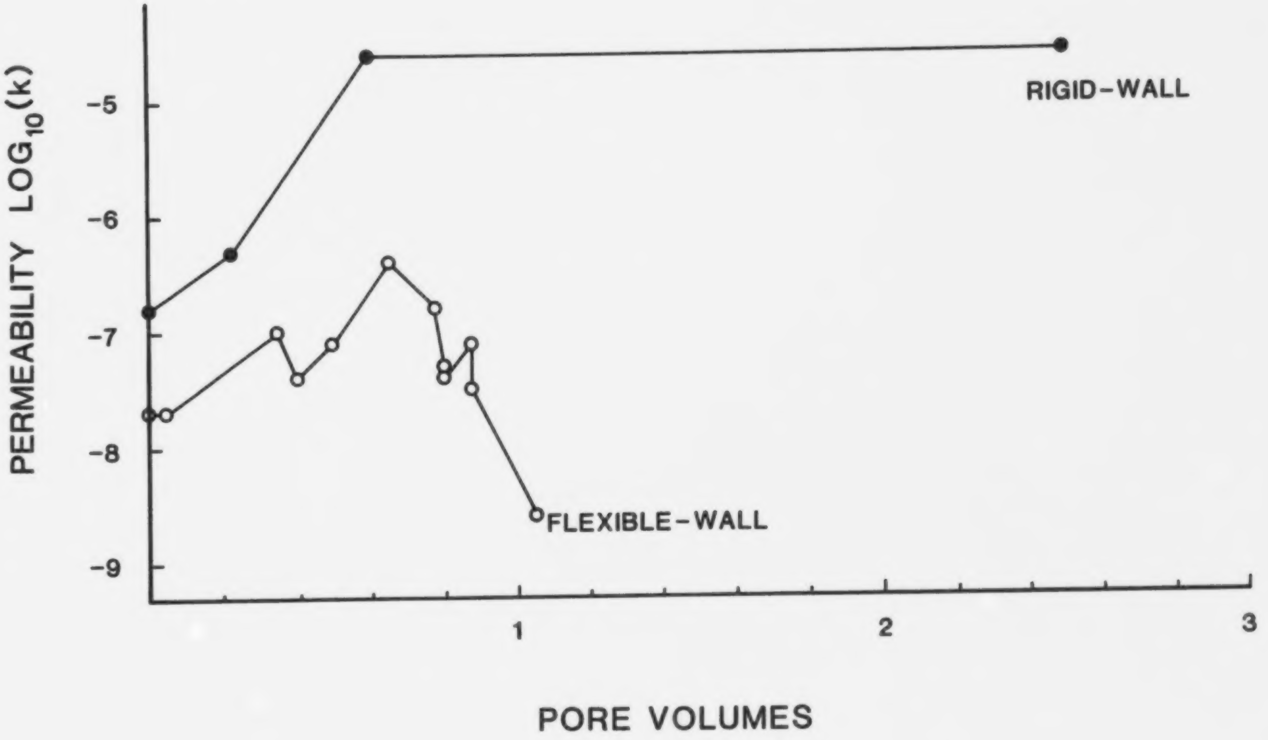
Figure 46. Kaolinite Clay with trichloroethylene in rigid-wall and flexible-wall permeameters

permeability of $10^{-4.86} = 1.4 \times 10^{-5}$ cm/sec after about 2.4 pore volumes of flow, and values appear to be still increasing at that point. The flexible-wall test exhibits immediately decreasing permeability. At termination (.08 pore volumes) the recorded permeability is $10^{-8.78} = 1.7 \times 10^{-9}$ cm/sec, roughly four orders of magnitude smaller than the final permeability of the rigid-wall test. In these tests, it was noted that the flexible-wall sample was pre-permeated with .01N $CaSO_4$, but the rigid-wall sample was not.

Similar results were obtained on a pair of tests in which Kaolinite Clay was permeated with heptane at a gradient of 200 (Figure 47). The rigid-wall sample permeability increased to $10^{-4.60}$ (2.5×10^{-5} cm/sec) over the first 0.6 pore volume, and the test then quickly passed an additional 2 pore volumes. At termination, the flexible-wall test, which was pre-permeated as in the case of the trichloroethylene, reached the level of $10^{-8.64}$ (2.3×10^{-9} cm/sec). It should however be noted that at an increased gradient of 300, five replicated flexible-wall tests on heptane gave extremely inconsistent results; in two cases the permeability decreased beyond 10^{-9} cm/sec, but in the remaining three the behavior was qualitatively similar to the rigid-wall test described above (see Figure 7). Two additional flexible-wall tests at a gradient of 100 both resulted in decreased permeabilities as in Figure 47.

Figure 48, taken from (20), illustrates typical plots of permeability versus pore volumes of flow obtained by Forman and Daniel for flexible-wall, rigid-wall and consolidation cell tests of Kaolinite Clay using 100% methanol as permeant (no 100% methanol tests on Kaolinite Clay using rigid-wall or consolidation cell permeameters were included in this data base). This figure implies that rigid-wall permeabilities should be about triple the corresponding flexible-wall permeabilities in this case, or a difference of about one-half order of magnitude. Likewise, rigid-wall tests conducted by Brown and Anderson on Hoytville Illite and Lufkin Noncalcareous Smectite with 100% methanol and heptane can be compared with tests on the same clay soils and permeants by Daniel. These tests are not matched in terms of the gradients used or the water content at which specimens were prepared as indicated in Figures 49-52. In addition, the characteristics of the soils as reported by the two labs are rather dissimilar, as will be discussed below. As shown in Figures 49-52, the behavior of the tests is consistent with earlier observations. For both clays, when heptane is the permeant, rigid-wall tests result in permeabilities increasing to on the order of 10^{-6} cm/sec or beyond by termination, whereas flexible-wall tests exhibit very small terminal permeabilities (Figures 49 and 50). When methanol is the permeant (Figures 51 and 52) the rigid-wall tests again yield permeabilities of 10^{-6} cm/sec or more, but the flexible-wall tests equilibriate at levels several orders of magnitude lower. The flexible-wall tests with methanol were run at gradients ranging from 10 to 300, and a considerable amount of the variation observed among the levels they achieve appears to be explained by gradient differences (as well as molding water content), a point which is discussed in Section 4.5.

In summary, although there are very few well-matched comparisons of flexible-wall and rigid-wall tests using 100% concentrated chemicals in the data set, the comparisons that are available give no reason to expect that their behavior should be at all similar. This is not surprising, since the



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 47. Permeability of Kaolinite Clay with heptane in rigid-wall and flexible-wall permeameters

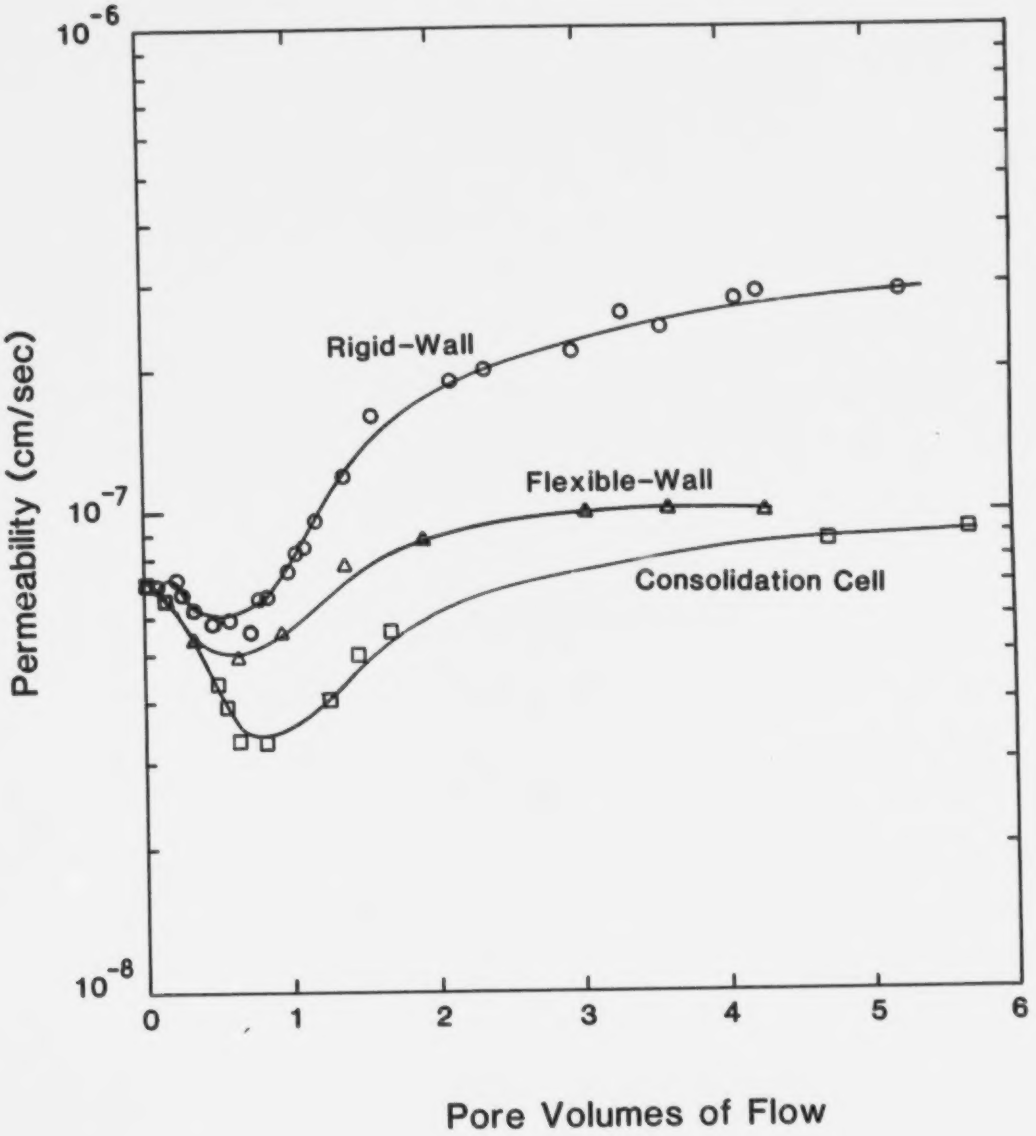
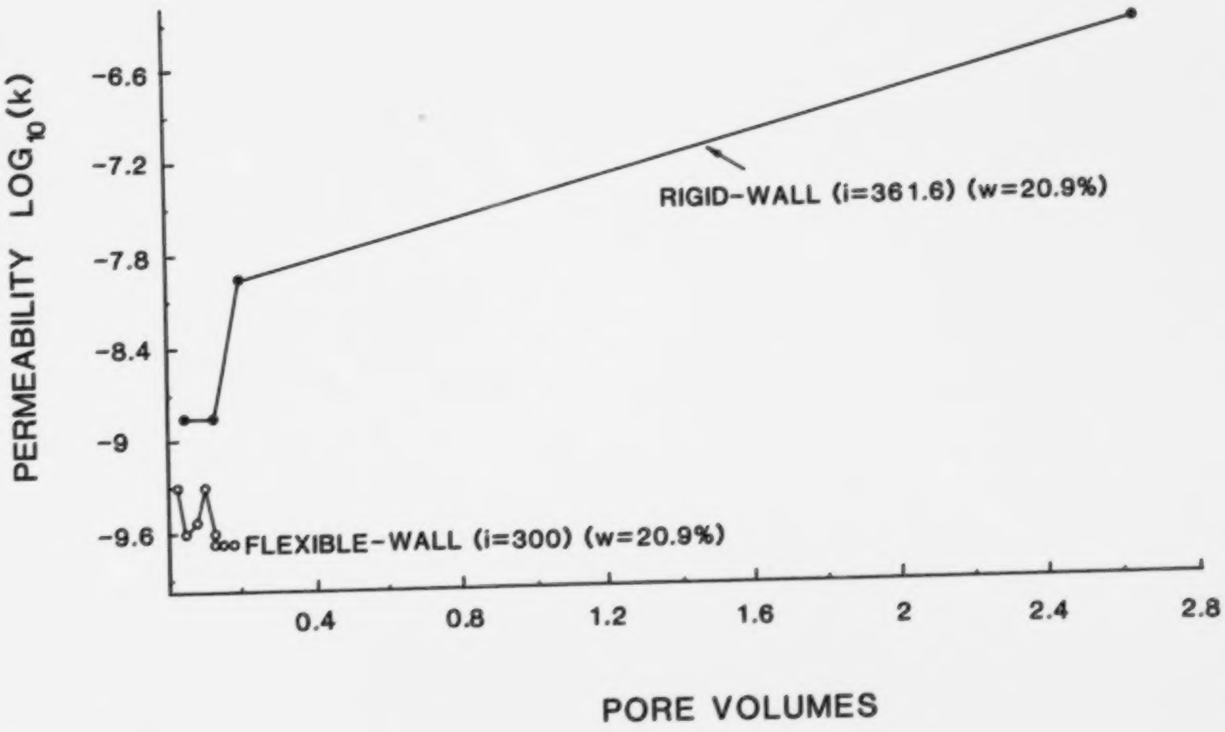
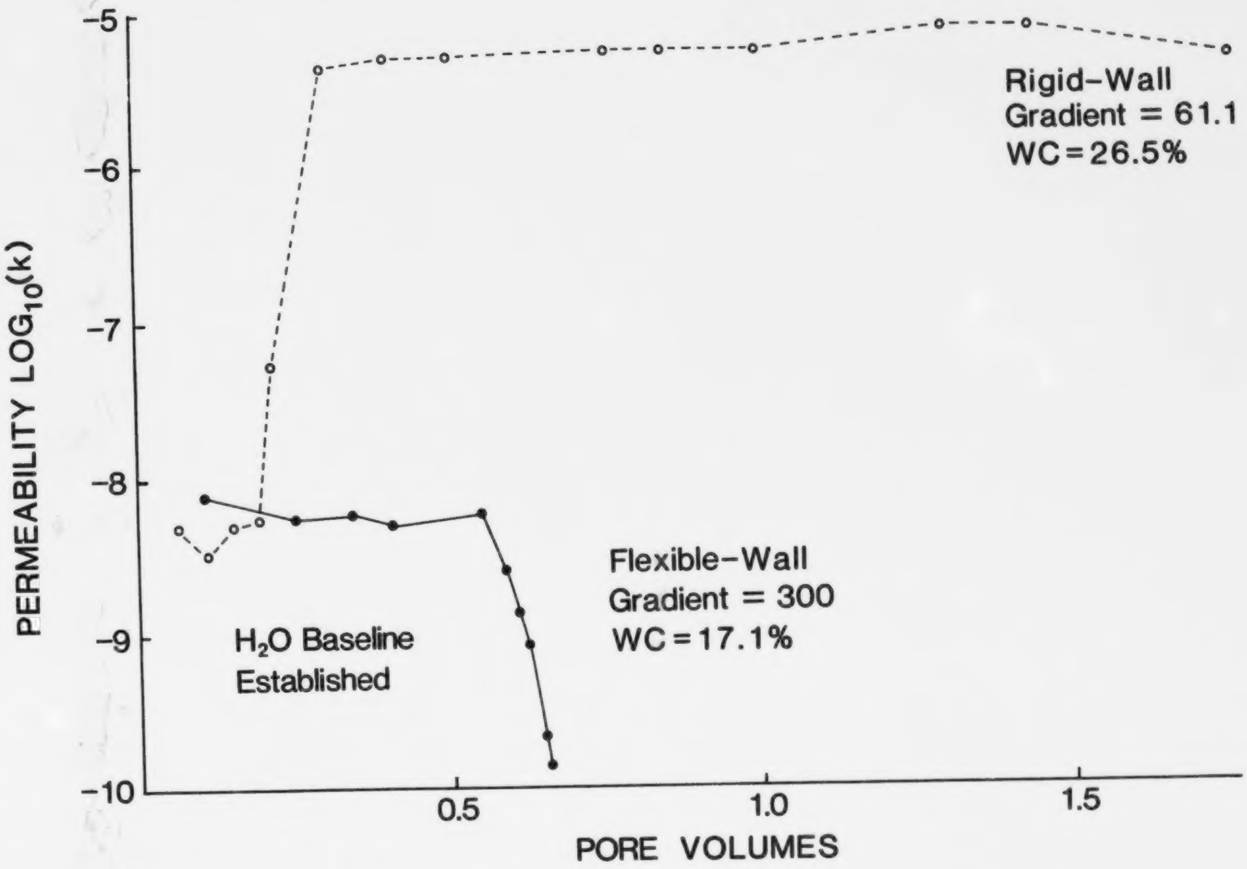


Figure 48. Permeability versus number of pore volumes of flow for Kaolinite Clay permeated with methanol with gradients of 250 or 300 (Daniel, 1984)



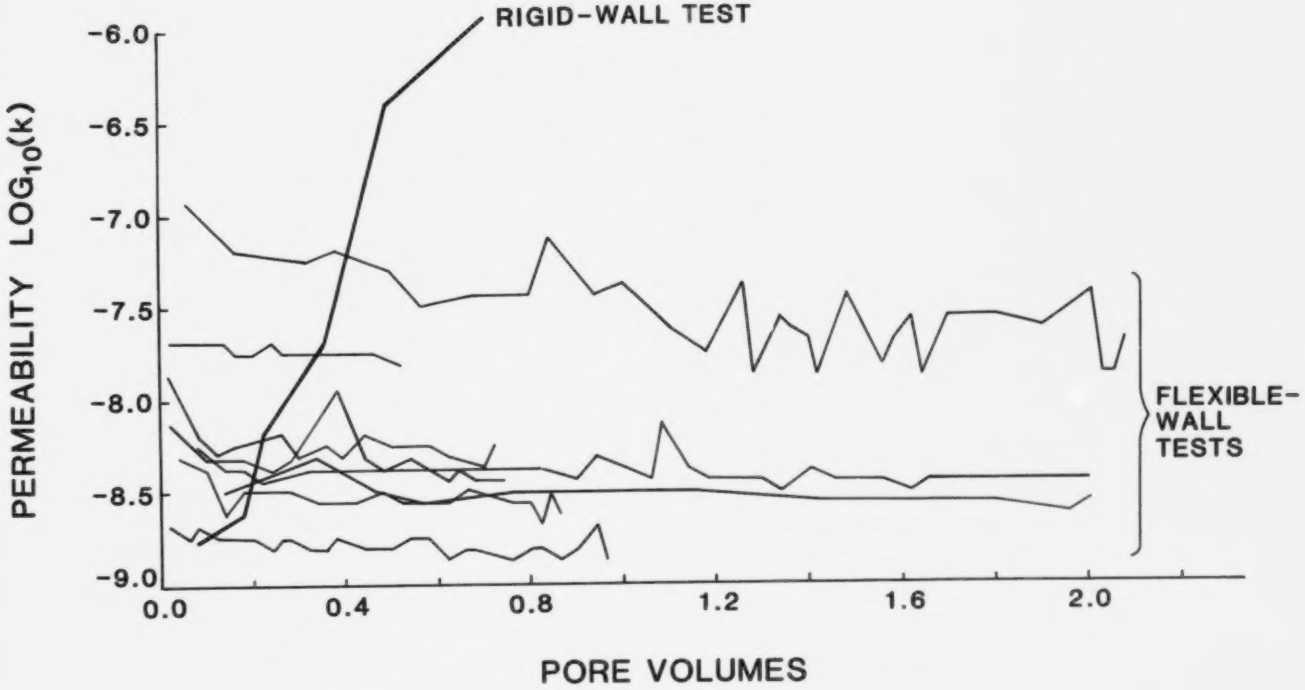
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 49. Permeability of Lufkin Noncalcareous Smectite with heptane in rigid-wall (Brown and Anderson) and flexible-wall (Daniel) permeameters



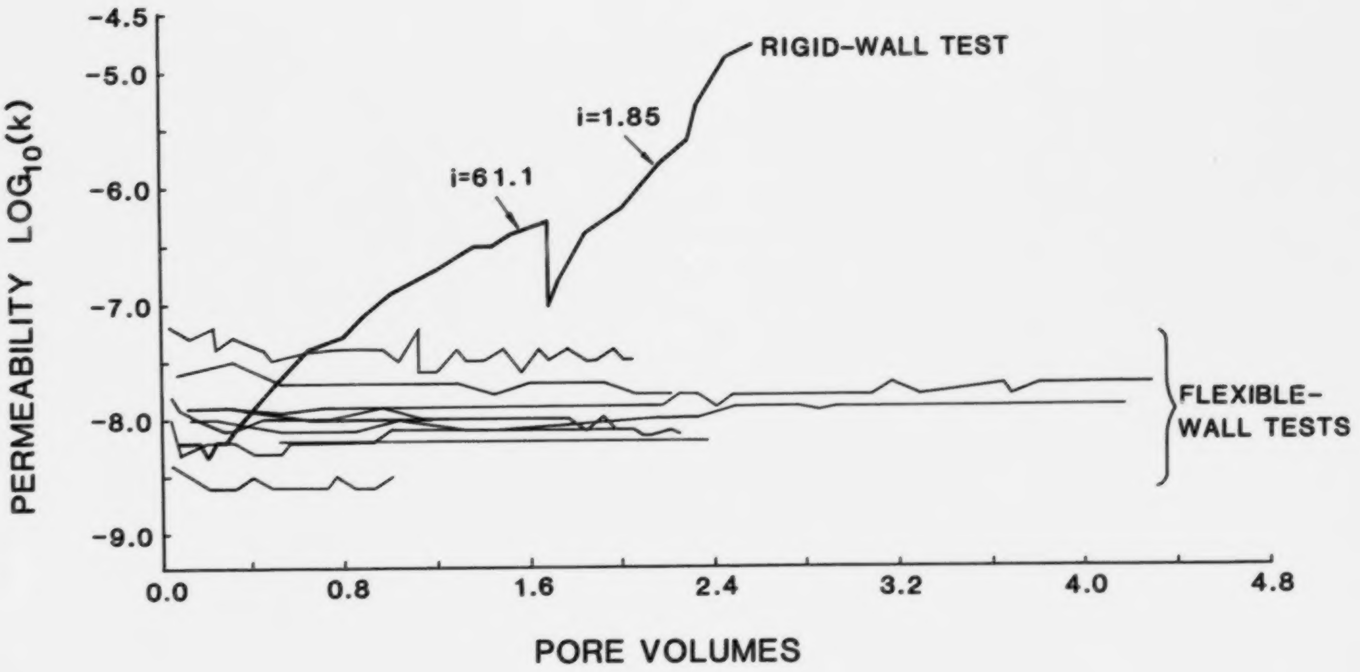
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 50. Permeability of Hoytville Illite with heptane in rigid-wall (Brown and Anderson) and flexible-wall (Daniel) permeameters.



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 51. Permeability of Lufkin Noncalcareous Smectite with methanol in rigid-wall (Brown and Anderson) and flexible-wall (Daniel) permeameters



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 52. Permeability of Hoytville Illite with methanol in rigid-wall (Brown and Anderson) and flexible-wall (Daniel) permeameters

flexible-wall encasement will conform to shrinkage that may be caused by chemical action of the permeants on the clay soils, whereas the rigid-wall permeameter will not.

In addition to the full-strength permeant tests described above, Daniel also conducted two series of tests on Kaolinite Clay soil at a gradient of 200 using both methanol and acetic acid at various concentrations in distilled water. For both rigid-wall and flexible-wall tests, concentrations of 20%, 40%, 60% and 80% were investigated.

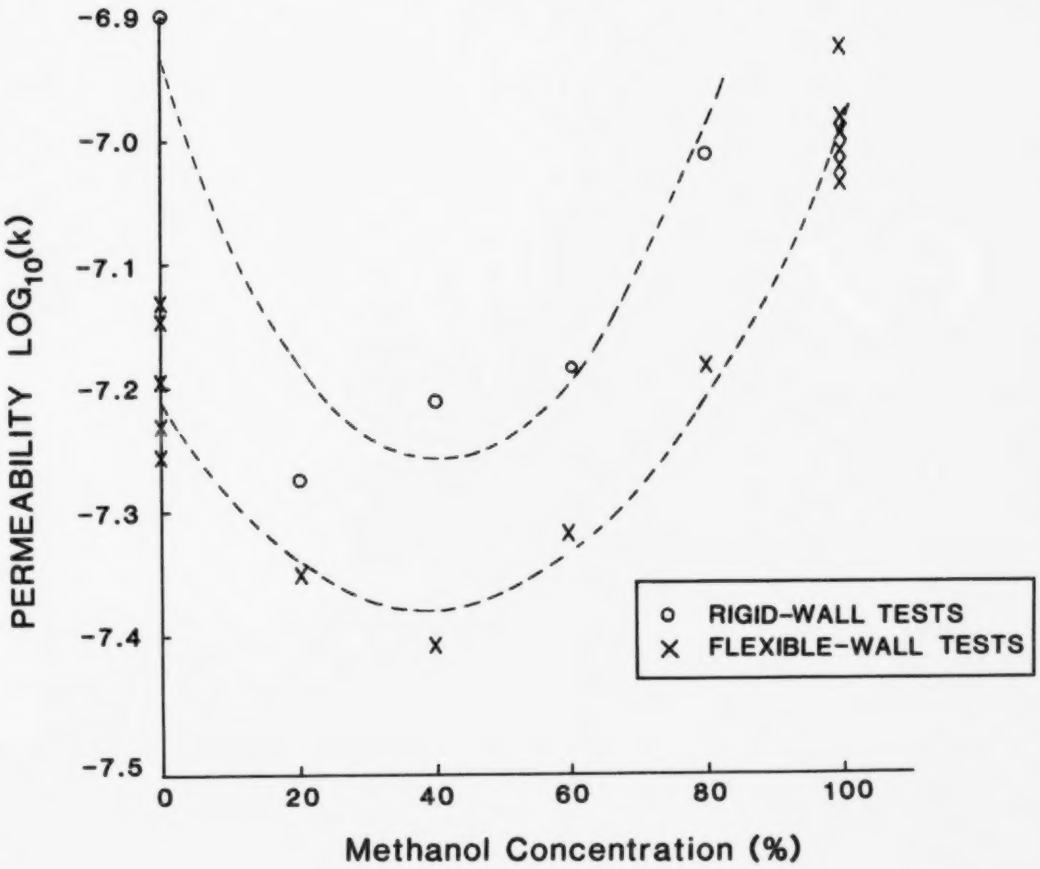
Figure 53 shows the final permeabilities of the rigid-wall and flexible-wall tests in the methanol series. For each test, these permeabilities were based on the total flow over an interval beginning with the last reading prior to one-half pore volume before termination.

Also shown in Figure 53 are the results of a rigid-wall test of Kaolinite Clay at a gradient of 200 using distilled water as permeant, which provides a baseline measurement of 0% concentration, and twelve flexible-wall tests, six using 100% methanol and six employing standard aqueous leachate (.01N $CaSO_4$). While the .01N $CaSO_4$ tests are not precisely equivalent to tests with distilled water, they seem to fit into the general pattern of data from the other flexible-wall tests, both visually and also in terms of statistical tests for lack-of-fit. The results of the analysis discussed below turn out to be essentially unaltered if these six tests are deleted from the data under consideration.

The twelve flexible-wall tests at 0% and 100% concentration also differ from the others in that they were run at gradients of either 100 (three at 0% concentration and three at 100%) or 300. However, as noted in Section 4.5, results of Daniel on Kaolinite Clay using either standard aqueous leachate or concentrated methanol were quite insensitive to gradient within this range, so that this factor is reasonably ignored here.

Figure 53 gives the rather surprising indication that small amounts of methanol may modestly decrease the permeability of this Kaolinite Clay, with the relationship between permeability and concentration turning upward as concentration approaches 100%. Fitting log permeability to a quadratic model in concentration for each permeameter type yields strongly significant tests for permeameter effect ($p = .0001$) and quadratic trend ($p = .0001$). There is no strong evidence that the two curves are not parallel ($p = .12$) and no evidence of model lack-of-fit ($p > .25$). A 95% confidence interval for the effect of the rigid-wall permeameter relative to the flexible-wall type is, on the log scale, (.13, .27), or an estimated percentage increase between $10^{.13}-1 = 35\%$ and $10^{.27}-1 = 86\%$. If the six .01N $CaSO_4$ tests are not used in the analysis, the interval becomes (.10, .27), corresponding to a percentage increase between 26% and 86%.

Even though there was no strong evidence that the realtive effect of permeameter type varied for concentrations between 0% and 80%, it would be unwise to attempt to extrapolate this to the 100% level. In fact, the test results of Daniel cited earlier imply a much more pronounced effect in that case.



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 53. Permeabilities of Kaolinite Clay in flexible-wall and rigid-wall permeameters at various concentrations of methanol

Figure 53 may show somewhat more variability among the rigid-wall tests than among the flexible-wall tests. This is consistent with Figures 54 and 55, which show log-permeability versus pore volumes of flow for the rigid-wall and flexible-wall tests at 20%, 40%, 60% and 80% concentrations of methanol. Comparison of these figures, which are drawn to the same scale, shows that the rigid-wall tests are somewhat more erratic than the tests on flexible-wall samples.

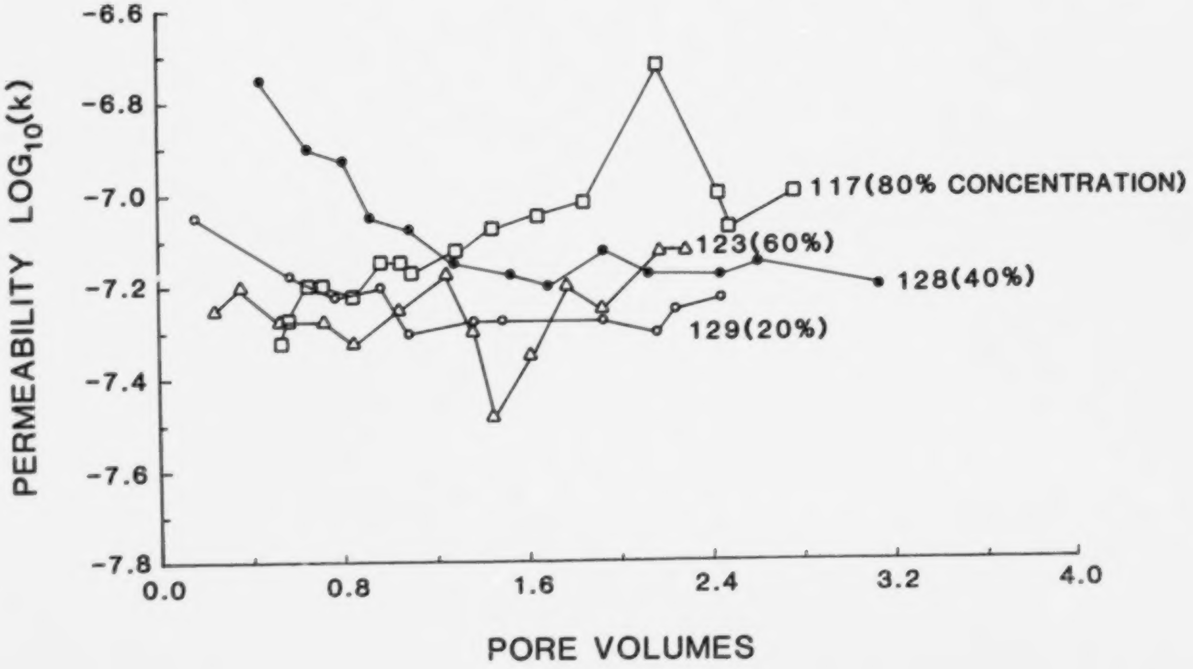
Figure 56 shows permeabilities at termination for the tests in the acetic acid series. The behavior of the flexible-wall tests is very similar to those in the methanol series. The rigid-wall test results are quite scattered and include one test at 20% which essentially failed at about 3.5 pore volumes, after which its permeability rose quickly beyond 10^{-4} cm/sec. The rigid-wall tests at 20% and 40% have not equilibrated, and the stability of the one at 100% is questionable. Therefore it is difficult to make any reasonable quantitative estimate of the relative effect of permeameter type for this series.

Figures 57 and 58 are pore volume plots for the rigid-wall and flexible-wall tests respectively, and are drawn to the same scale (much different than the scale of Figures 54 and 55). At this scale, little detail can be distinguished concerning the flexible-wall tests, but the figures graphically illustrate the fact that in this series the rigid-wall tests behaved far more erratically than did the tests on flexible-wall samples.

Peirce, in his investigation of the effects of dilute solutions of $FeCl_3$ and $Ni(NO_3)_2$ on Faceville Clay, Hoytville Illite and Whitestore Montmorillonite soils, included several flexible-wall tests in addition to the rigid-wall tests that were primarily employed. Plots of these test results are shown in Figures 59, 60, and 61, and a more detailed description of the Faceville series is given in Section 4.4.

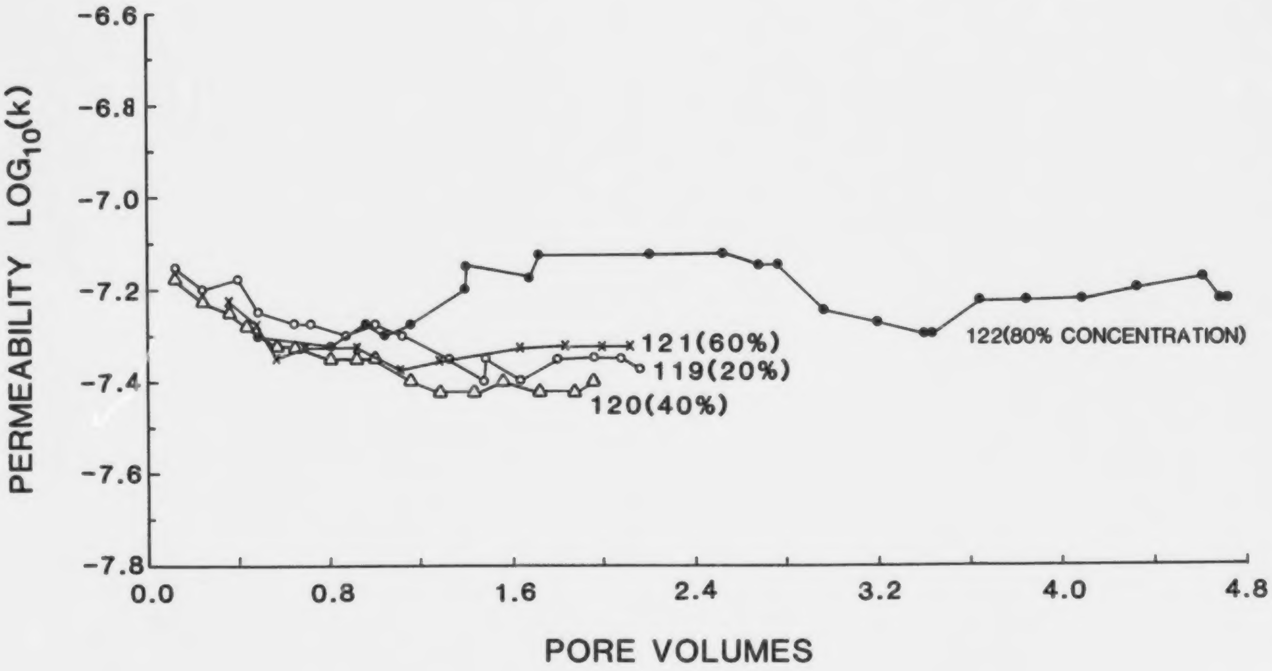
At first glance, Figures 59 and 60 appear to give the impression that the flexible-wall apparatus yields slightly lower permeabilities than does the rigid-wall permeameter, if the tests are run long enough to overcome the initially higher flexible-wall permeabilities (which were vacuumed). Figure 61, which depicts the Hoytville Illite series, could be possibly interpreted in the same way, if one were to accept the hypothesis that the tests are continuing a very slow approach to equilibrium at the time they are terminated. However these tests were never intended to form the basis of a statistical comparison of permeameter types, and for each of the three clays, the comparison between permeameter types has been confounded with other factors to which this minor effect could reasonably be attributed. In each case, the rigid-wall tests were moisturized in a common batch and tested concurrently; the flexible-wall tests were prepared separately from the rigid-wall tests by a second technician, and were compacted to a somewhat greater initial density. Average gradients of these (falling head) tests were not closely matched across permeameter types. Therefore, any one of several factors is sufficient to explain the minor differences observed in the final permeabilities, so that no unambiguous comparisons can be made.

As discussed in Section 4.4, four of the twelve rigid-wall tests on Faceville Clay (Figure 59) were control tests using only standard aqueous



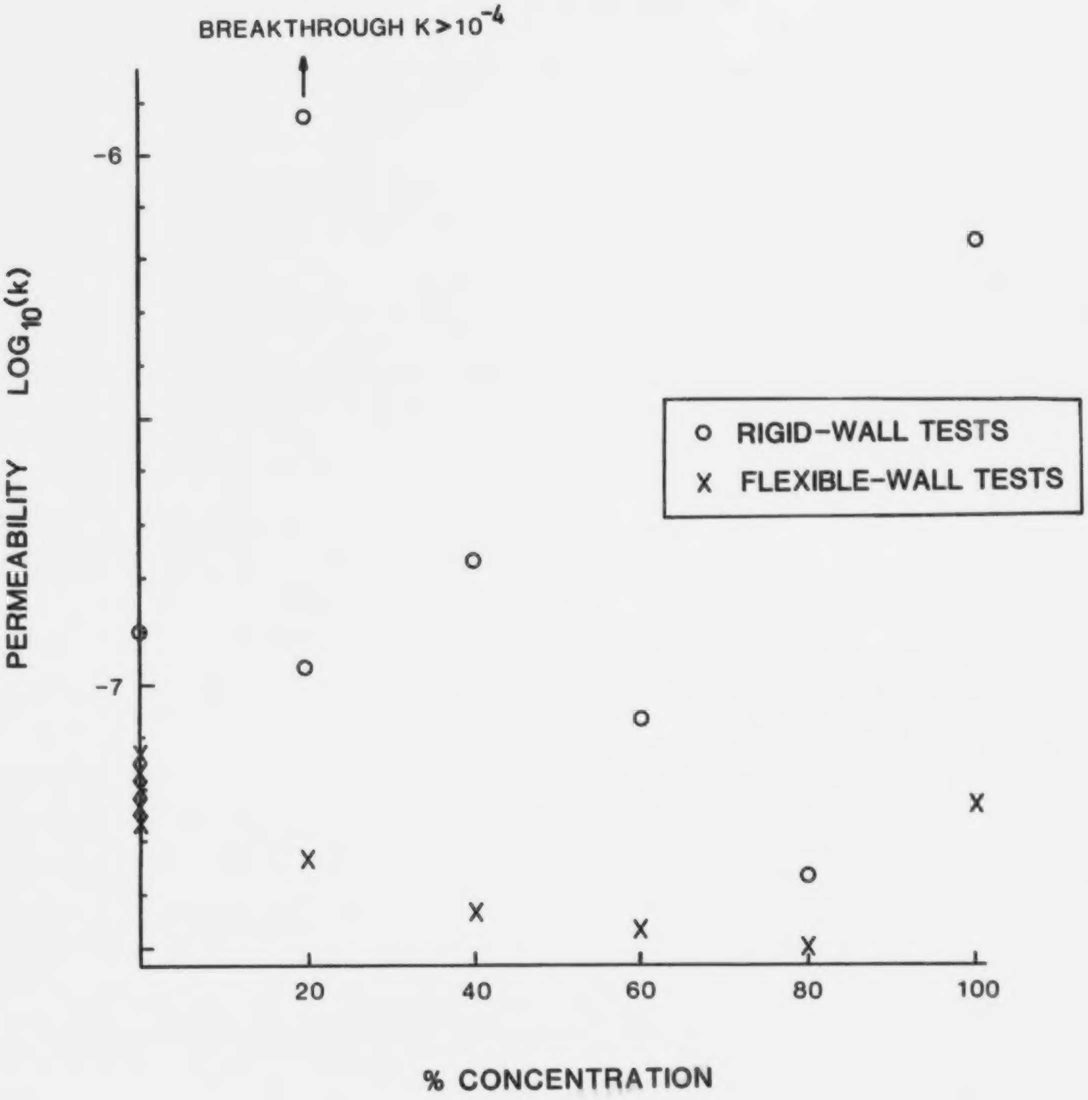
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 54. Permeability of Kaolinite Clay with methanol in rigid-wall permeameters with i = 200



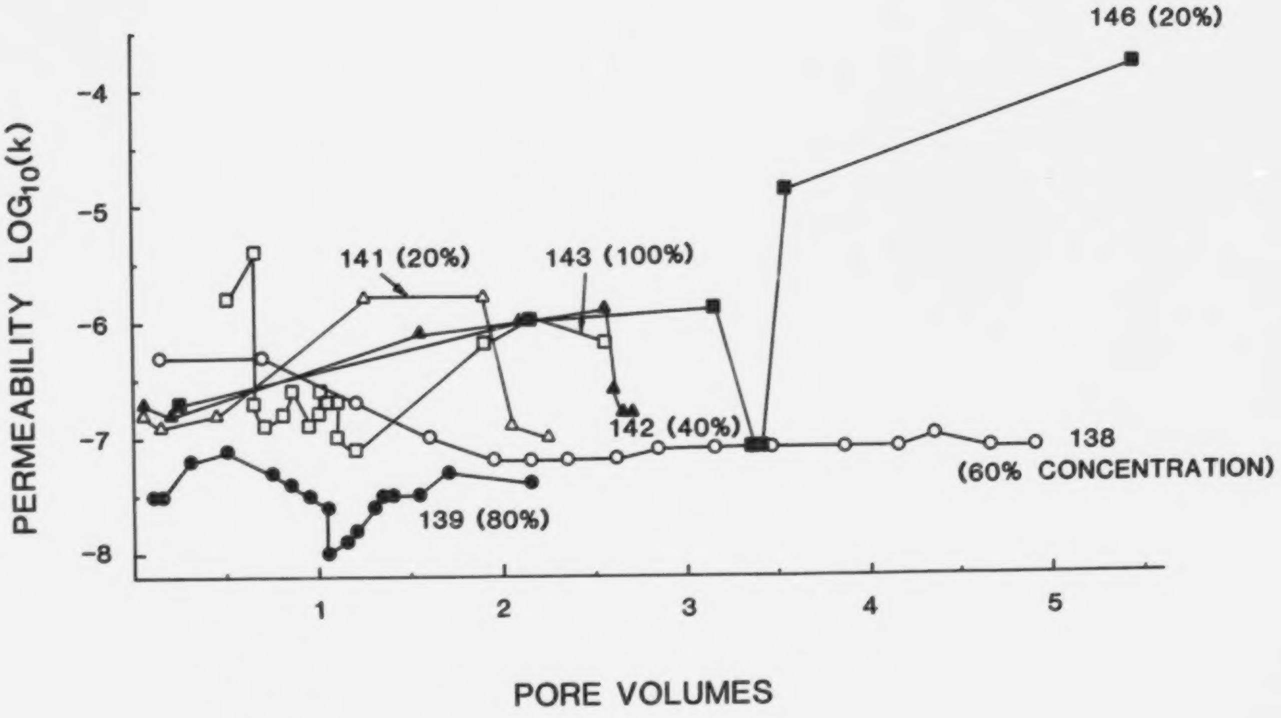
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 55. Kaolinite Clay with methanol in flexible-wall permeameters with i = 200



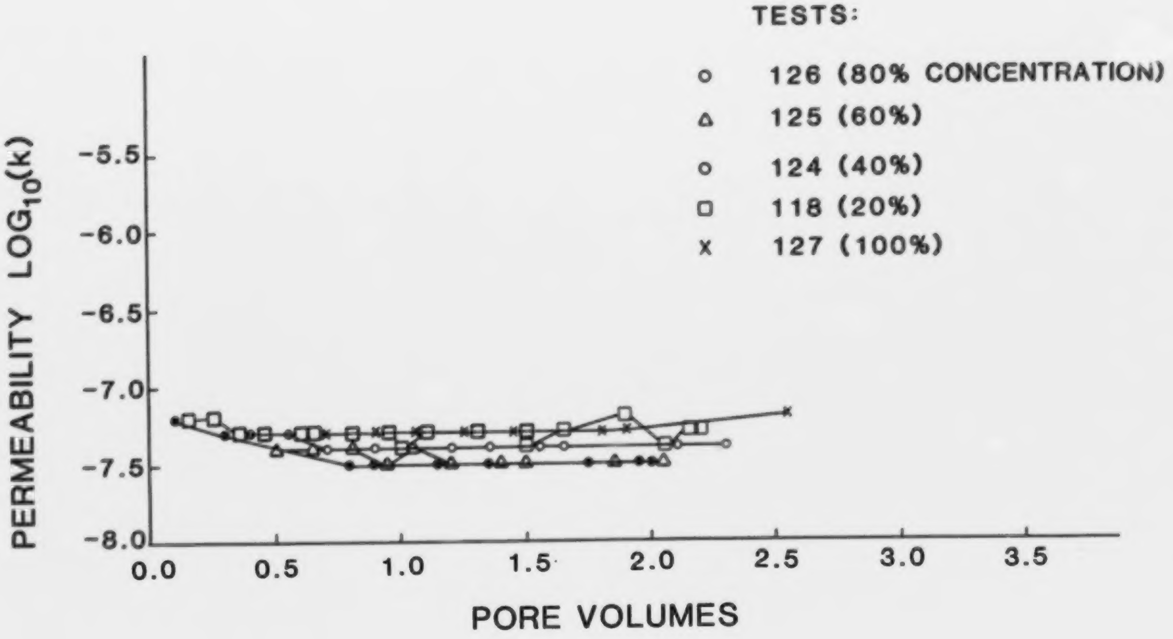
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 56. Permeabilities of Kaolinite Clay in rigid-wall and flexible-wall permeameters at various concentrations of acetic acid



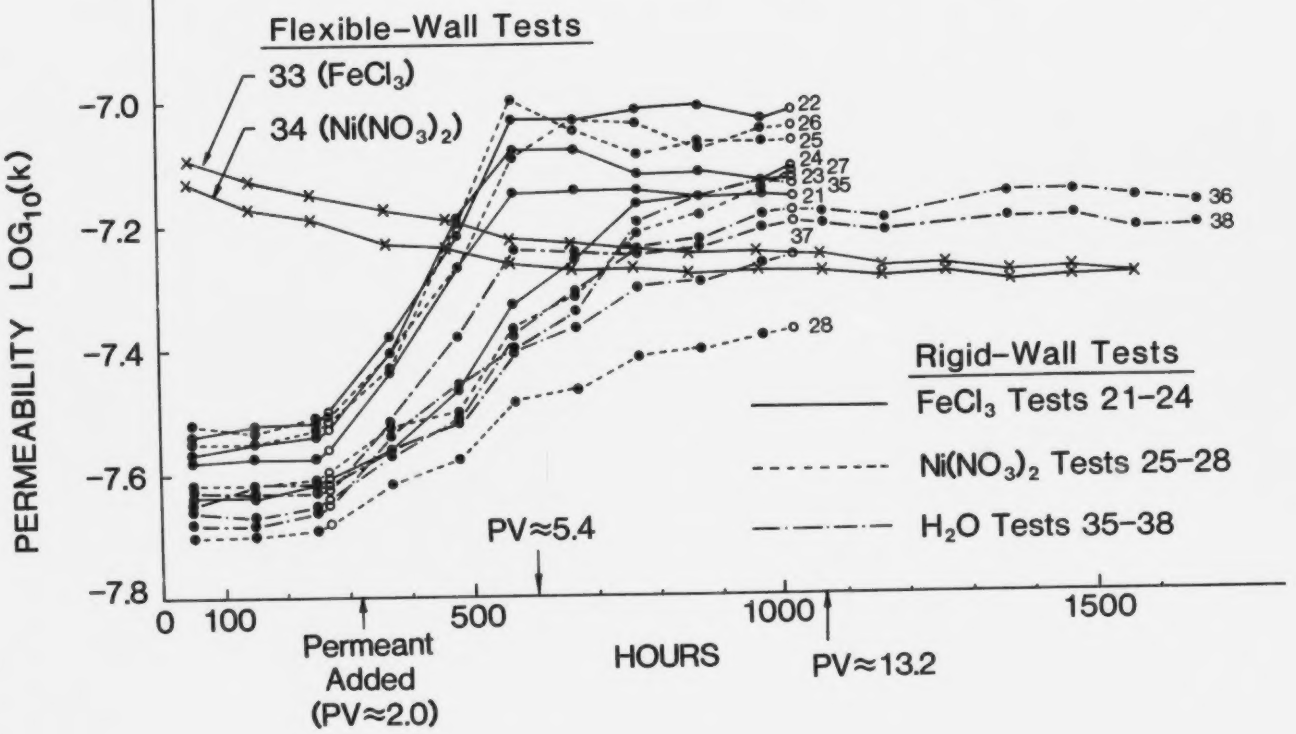
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 57. Kaolinite Clay with acetic acid in rigid-wall permeameters with i = 200



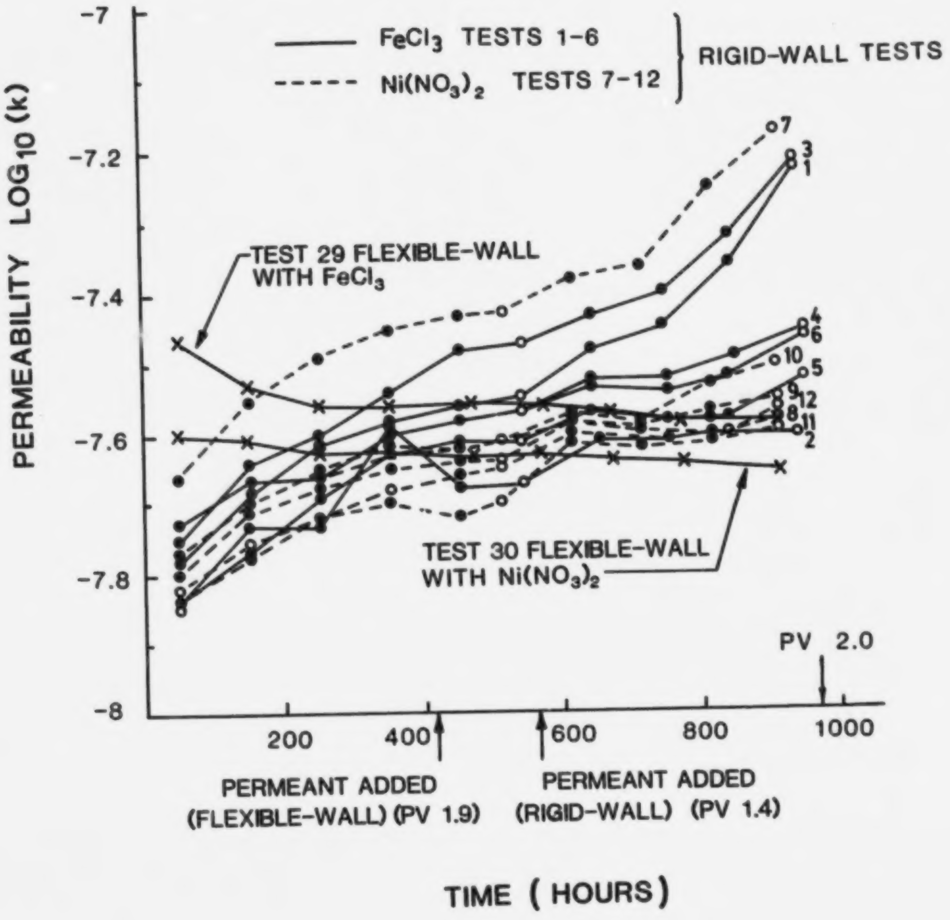
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 58. Kaolinite Clay with acetic acid in flexible-wall permeameters with i = 200



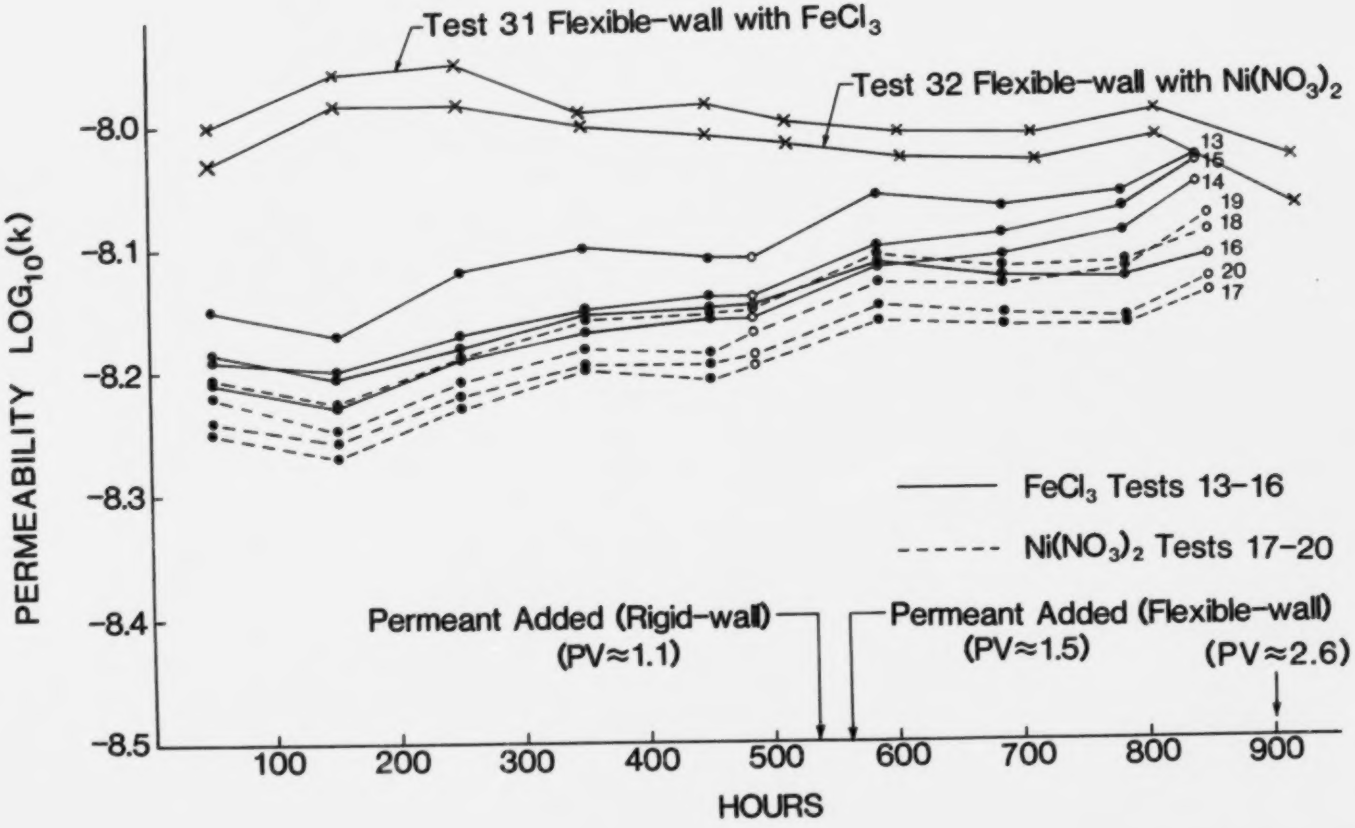
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 59. Permeability of Faceville Clay (Comparison of rigid-wall and flexible-wall tests)



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 60. Permeability of Whitestore Montmorillonite (Comparison of rigid-wall and flexible-wall tests)



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 61. Permeability of Hoytville Clay (Comparison of rigid-wall and flexible-wall tests)

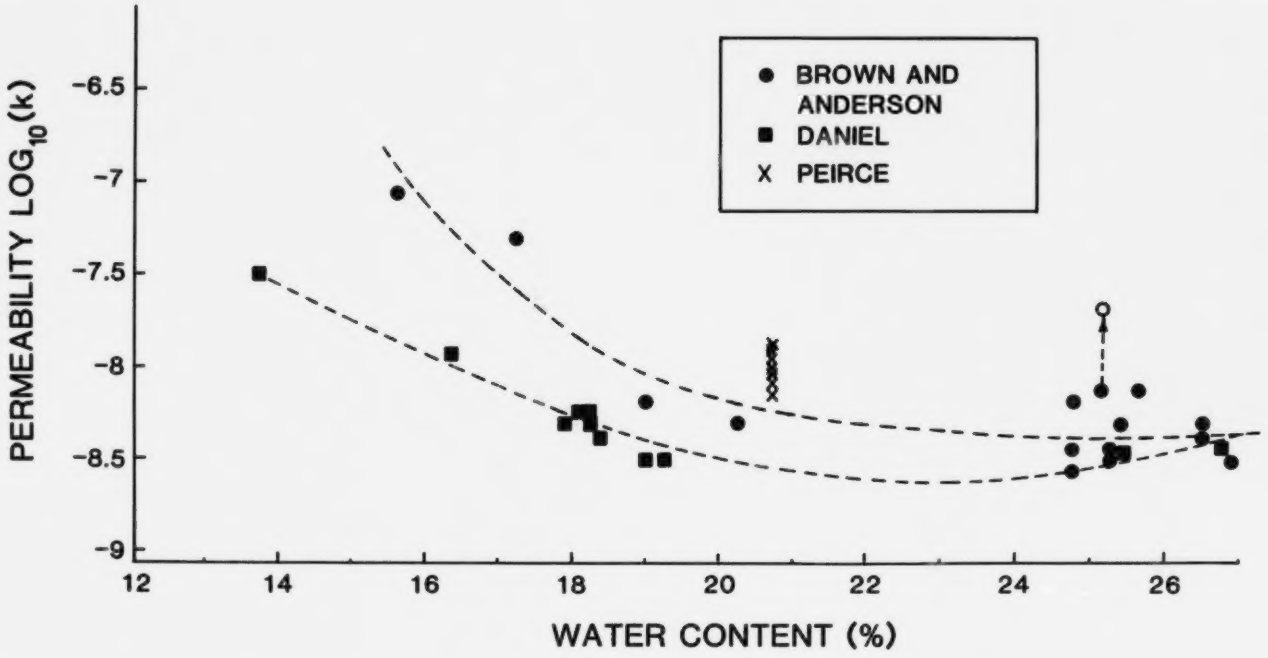
leachate. The behavior of these tests was quite similar to that of the tests which had received (very dilute) solutions of $FeCl_3$ and $Ni(NO_3)_2$, and in fact statistical tests indicate no effect of these permeants relative to that of standard aquaeous leachate. Thus the nearly one-half order of magnitude increase in permeability observed after the introduction of chemical permeants can not be attributed to the permeants themselves, but may be caused by a failure to have achieved equilibrium during the baseline period, or possibly to an environmental effect whose magnitude is far greater than could be accounted for by the effect of temperature on viscosity. Neither do the flexible-wall tests on Faceville Clay indicate any evident effect due to permeant, and their approach to equilibrium is very consistent. In varying degrees, each of Figures 59, 60 and 61 seems to suggest that flexible-wall samples may equilibrate more readily than do rigid-wall samples when compared under reasonably similar circumstances. This observation is consistent with the relative behavior of the two types of permeameters in Daniel's methanol and acetic acid series.

Shown in Figure 62 are results of water permeation tests of Hoytville Illite, by Brown and Anderson, Daniel and Peirce. The data from Brown and Anderson come from initial water permeation of tests to which chemicals were subsequently added, making up eight of the points at water contents near 25%, and seven additional preliminary tests which were conducted prior to the main experiment to estimate the sensitivity of the soil to molding water content. As discussed in Section 4.1 one of the tests was used as a water control test during the main experiment and its permeability increased by about .4 of an order of magnitude over 2.66 pore volumes, as shown in Figure 62. Peirce's data also came from preliminary water permeation, as shown in Figure 61.

Direct comparison of the results from different researchers is difficult, because they typically did the bulk of their experimentation at different water contents, and the distribution of water contents investigated at each lab is not suitable to realistically estimate the shape of the water content permeability relationship. The curves drawn in Figure 62 were obtained by least squares by fitting log permeability to separate cubic models in water content. The most noticeable feature of the plot is that Daniel's and Brown and Anderson's results at the very low water contents are somewhat dissimilar, differing by roughly .8 of an order of magnitude. However, no inference can be drawn; not only were the tests run in different permeameter types, but also they were run at different gradients, and more importantly, the homogeneity of the soils used at the two labs is questionable. Brown reported percentages of clay, silt and sand of 47%, 38% and 15% for his soil. The corresponding figures for Daniel were 65%, 22% and 13%, and for Peirce were 47%, 39% and 14%.

4.7 COMMENTS ON EQUILIBRATION WITH CHEMICALS IN RIGID-WALL PERMEAMETERS

Rigid-wall tests, which were performed on replicated specimens of clay soils using the same chemical permeants and gradients, were provided by Peirce and by Brown. Peirce's tests have been discussed in previous Sections.



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 62. Effect of molding water content on the permeability of Hoytville Illite (with .01N $CaSO_4$)

Brown's tests, which are discussed in this Section, made use of Elgin Kaolinite Clay, Micaceous Clay and Bentonite Clay and used concentrated acetone and xylene as permeants. These chemicals generally caused very considerable increases in permeability over the water baselines.

Altogether, results of 65 tests with chemical permeants were provided. Thirty-three of these involved presaturation with .01N $CaSO_4$ accomplished by permeating generally on the order of one pore volume of standard aquaeous leachate prior to introducing chemicals. The remaining 32 tests were not saturated in this manner. Tests on Elgin Kaolinite Clay and Micaceous Clay were performed at gradients of 31, 91 and 181, while those on Bentonite Clay were run at gradients of 91, 181 and 272. In all but five cases at least one pore volume of chemical was passed, and in some cases as much as six pore volumes were passed. All tests used the constant head method.

Results of these tests were reasonably consistent qualitatively. In 24 of the 33 presaturated tests (73%), permeabilities increased by one order of magnitude or more over their water baselines by the time the tests were discontinued. In most cases the increases were very substantial, often in the range of two to four orders of magnitude. An additional two tests increased nearly one order of magnitude and appeared almost certain to have increased further had they been continued. Thus in about 79% of the presaturated tests, consistent and meaningful increases in permeability took place. Similarly, 24 of the 32 nonsaturated tests (75%) exceeded the average baselines established in water permeation of the same clays by at least one order of magnitude, and usually much more than this amount. Three others, which were increasing in permeability, appeared likely to have exceeded this level if they had not been terminated.

With regard to summarizing or describing the variability of these data, the writers do not feel that a numerical measure such as standard deviations of the terminal permeabilities obtained in the tests provides a meaningful description. Many of the tests have clearly not achieved an equilibrium at termination, and for others any equilibration which may appear to have taken place is questionable. It is emphatically not the intention of this last statement to imply that rigid-wall tests using such chemicals as acetone and xylene are not meaningful or do not provide useful data. On the contrary, the selection of equipment and methods in permeability testing must ultimately depend on the degree to which the experimenter believes the test to adequately model the environment and phenomena he or she is investigating. In some circumstances, however, it may be that the notion of an equilibrium permeability, which is obtained after adding chemical permeants such as those used in these experiments, may not be the most useful way to look at the data, since the effects of these chemicals is to promote the development of shrinkage cracks and channeling in the specimens, leading to abrupt and often very dramatic increases in permeability. In some ways, the problem is more one of an analysis of the time to failure of the specimens, where the term "failure" must be defined in some sense which is appropriate to the investigation at hand. If one takes this point of view, then a description of the variability of such data based on measures such as standard deviations of terminal permeabilities has the tendency of making the data appear more inconsistent than is actually the case.

Graphical methods for summarizing and describing data such as those discussed in this Section are very useful. In circumstances in which permeability changes radically over time, it was found that the use of the usual plots of permeability versus pore volumes of flow can sometimes give a misleading impression of the nature of the progression of the tests. On the other hand, plotting permeability against the actual time on test is sometimes also unsatisfactory; in cases where permeabilities increase by several orders of magnitude during the test, a substantial amount of flow can occur over time intervals of negligible length in comparison to the entire time on test.

One graphical method which has been found to be very useful is to simply provide both plots on a single page so that they may be viewed simultaneously; the time plot gives a good sense of the overall progression of the test, and the pore volume plot is useful for inspecting the "local" behavior of the test over short intervals of time in which the flow rate is very rapid.

Dual plots of this type are shown in Figures 63-74 for the twelve series of tests provided by Brown (each combination of one of three clay soils, two chemical permeants and presaturated versus nonsaturated tests). In each figure, the upper plot (A) is a time series plot of log permeability versus time on test, and the lower plot (B) graphs log permeability against pore volumes of flow. To increase the clarity of the plots, water permeation of the presaturated tests is not shown. The average water baseline for all presaturated tests of each particular clay is marked with an arrow on the vertical axis of each graph.

The raw data provided for these experiments did not include time at which measurements were recorded, but did include a listing of permeability versus pore volumes of flow. Time data were computed by applying Darcy's Law to each recorded measurement.

Let t_j be the time in days at which the jth measurement was recorded, with $t_0 = 0$. Let Q_j be the volume of flow over the jth time interval (cm^3) and k_j the average permeability over this interval (cm/sec). If i represents the gradient of the test, and A is the cross-sectional area of the specimen (cm^2) then from Darcy's Law

$$t_j - t_{j-1} = \frac{Q_j}{k_j iA \cdot 3600 \cdot 24} .$$

If ΔPV_j represents the change in cumulative pore volumes over the jth interval, e the void ratio of the specimen and L its length (cm), then

$$Q_j = \Delta PV_j eAL/(1+e).$$

Substituting this expression into the previous one and summing over the first n time intervals shows that the time t_n at which the nth measurement was recorded is given by

$$t_n = \left\{ \sum_{j=1}^{n} \frac{\Delta PV_j}{k_j} \right\} \cdot \frac{eL}{(1+e)i \cdot 3600 \cdot 24} .$$

Times computed by this formula should be considered to be only approximate, since they are sensitive to round-off in the recorded data. Also, the first permeability measurement in the raw data was usually made shortly after the test was set up, and so the recorded pore volumes of flow were generally only noted to be less than .1 . For this reason, the time origin is taken as the time at which the initial reading was made.

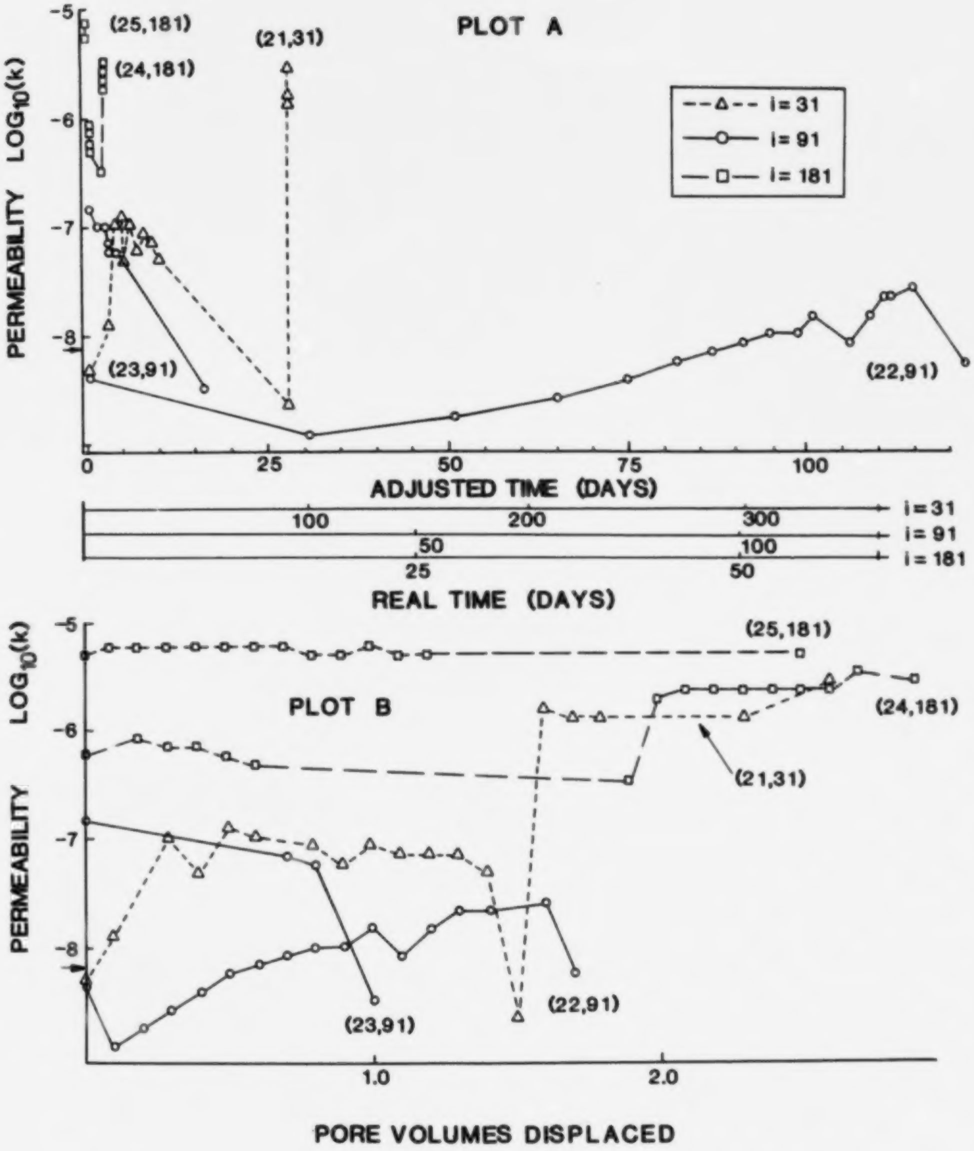
In Figures 63-A to 74-A. four time axes are given along the horizontal axis: one for tests run at each of the three gradients used in the experiments and a fourth axis labeled "adjusted time." This adjusted time is equal to be the actual time at which a recording was made times the gradient of the test divided by 100. Plotting permeabilities against this adjusted time in a sense compensated for the acceleration or deceleration of the tests due to increased or decreased gradient and is useful in overlaying plots of tests run at different gradients, since it tends to put them on the same scale. A gradient of 100 was arbitrarily chosen as a benchmark on which to base the computation of adjusted time.

Each test in Figures 63-74 is labeled with two numbers in parentheses. The first is a test number assigned for purposes of reference. The second number is the gradient at which the test was conducted.

It is convenient to begin discussion of the results by looking first at the presaturated Elgin Kaolinite Clay after the introduction of 100% acetone. Notice that the pore volume plot in Figure 64-B appears to indicate that the tests underwent a brief period of slightly decreased permeability due to swelling and then experienced gradually increased permeability which appears to be leveling off with the passage of time. In this plot, test 7 appears quite unusual. At about 3.6 pore volumes, its permeability shifts by about two and one-half orders of magnitude, stabilizing well above the level at which the remaining tests seem to be equilibrating.

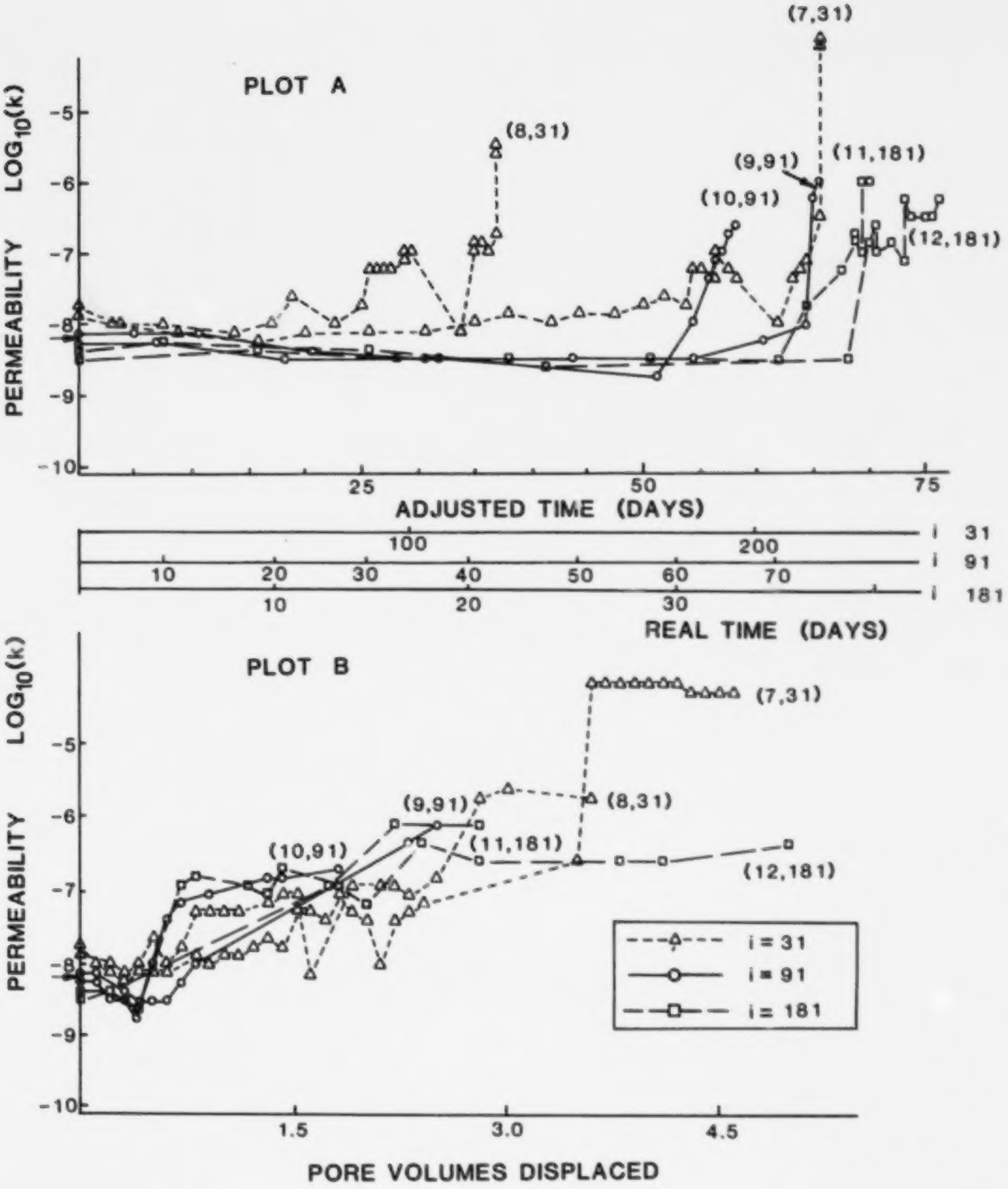
The time series plot in Figure 64-A gives a very different impression of the behavior of the tests. Here it is seen that the period of initially decreasing permeability actually spanned the major portion of the total time on test for most specimens. Then, permeabilities of each test abruptly increase. In each case an increase of at least two orders of magnitude is seen. These increases occur very episodically in real time. For example, test 9 remained below 10^{-8} cm/sec for over 70 days in real time, and then in about one day's time its permeability increased almost exactly two orders of magnitude, at which point the test was discontinued.

Similarly, test 7 remained at about 10^{-8} cm/sec for about 170 days. This was followed by fairly erratic behavior for about 30 days. From day 200 to day 211, permeability increased 3000% from 10^{-8} cm/sec to 3×10^{-7} cm/sec ($10^{-6.5}$) and then shot up to 10^{-4} cm/sec almost instantaneously. Even test 12, which according to the pore volume plot appears to be very stable, does not appear to be so when viewed in real time. For 38 days the permeability of this specimen remained reasonably constant between $10^{-8.3} = 5 \times 10^{-9}$ cm/sec and $10^{-8.6} = 2.5 \times 10^{-9}$ cm/sec. Then its permeability increased by a factor of 50 in 4 hours time. During the final four days it remained on test, its permeability gradually increased fivefold from $10^{-6.9} = 1.3 \times 10^{-7}$ cm/sec to $10^{-6.2} = 6.3 \times 10^{-7}$ cm/sec. The permeability versus pore volume plot of this



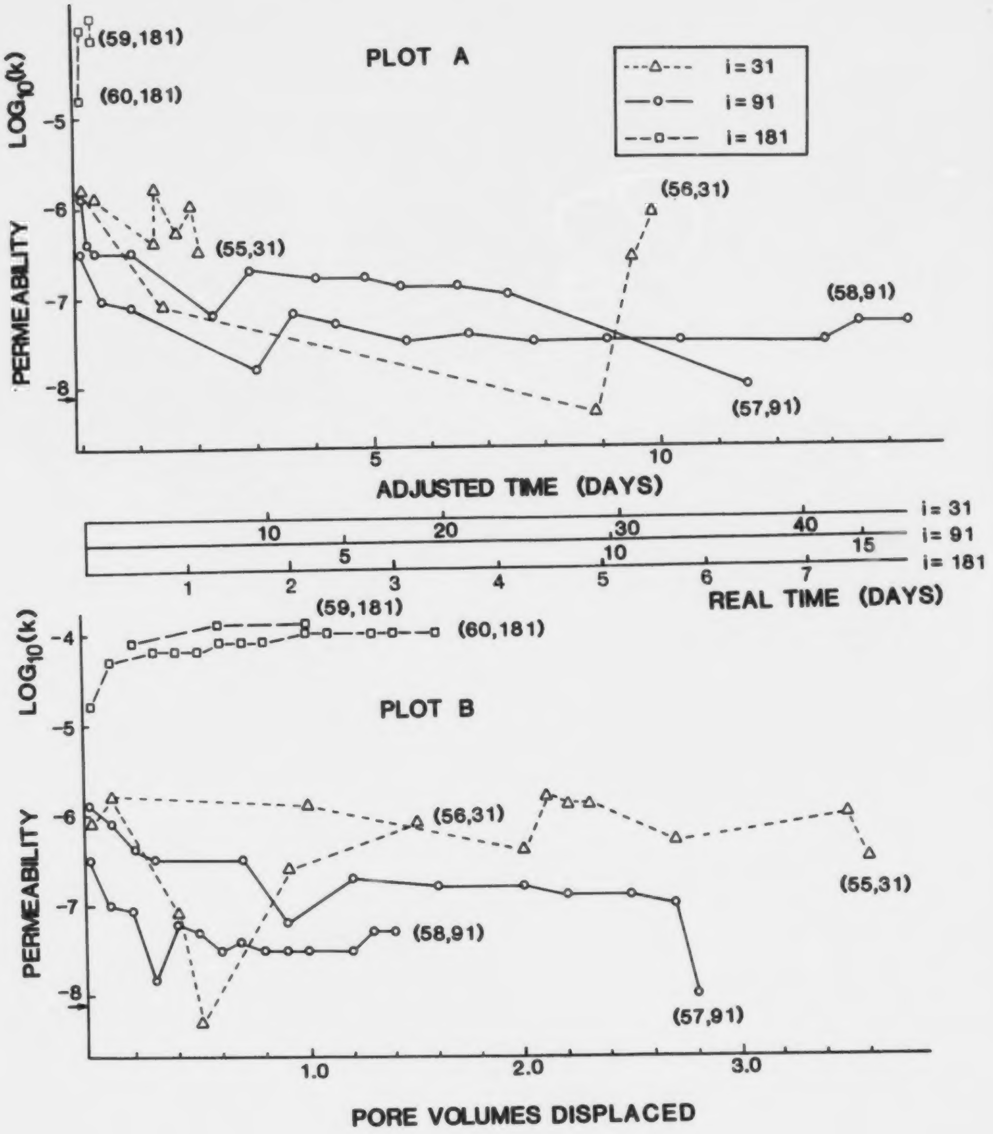
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figures 63-A and 63-B. Permeability of nonsaturated Elgin Kaolinite with acetone



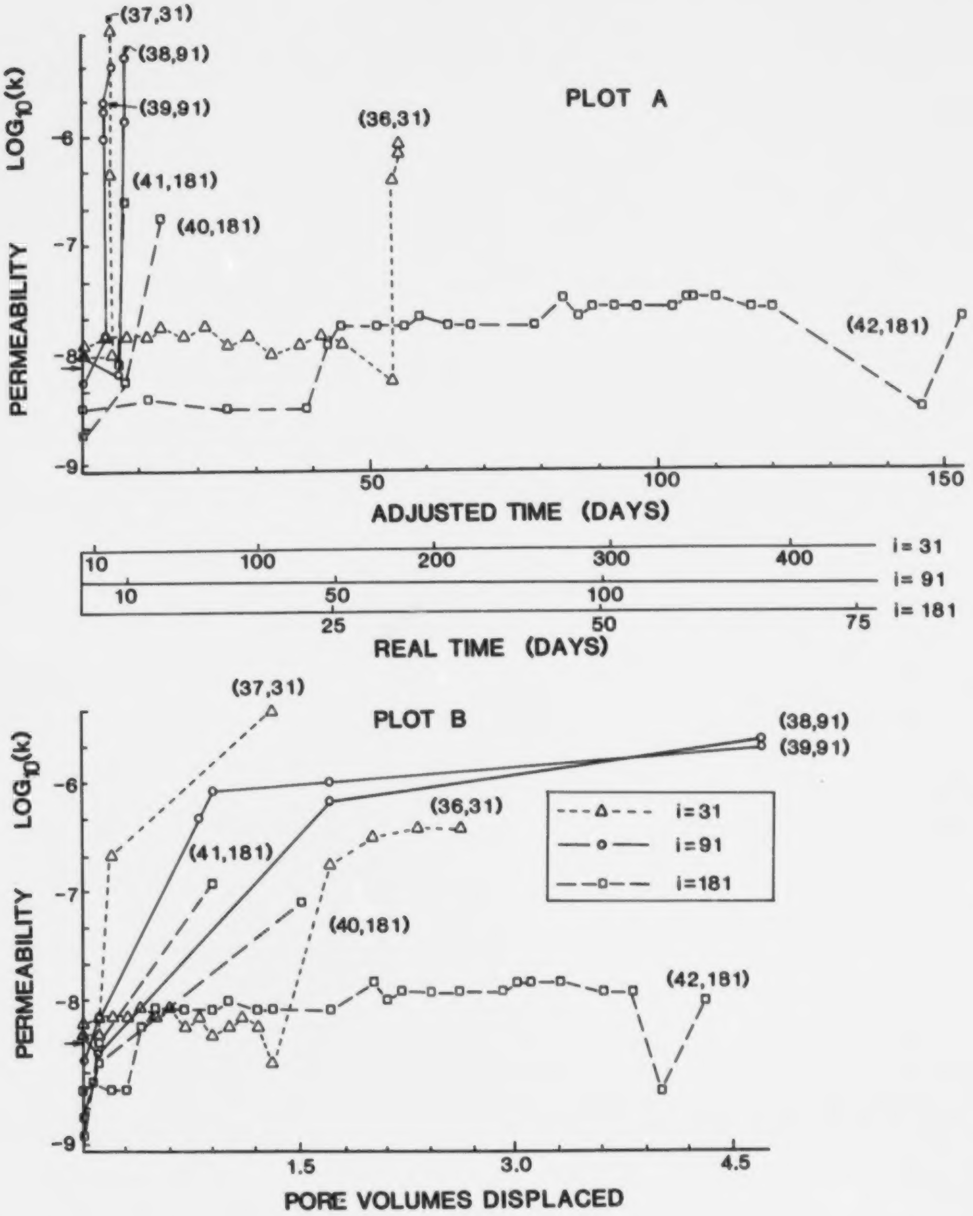
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figures 64-A and 64-B. Permeability of presaturated Elgin Kaolinite with acetone



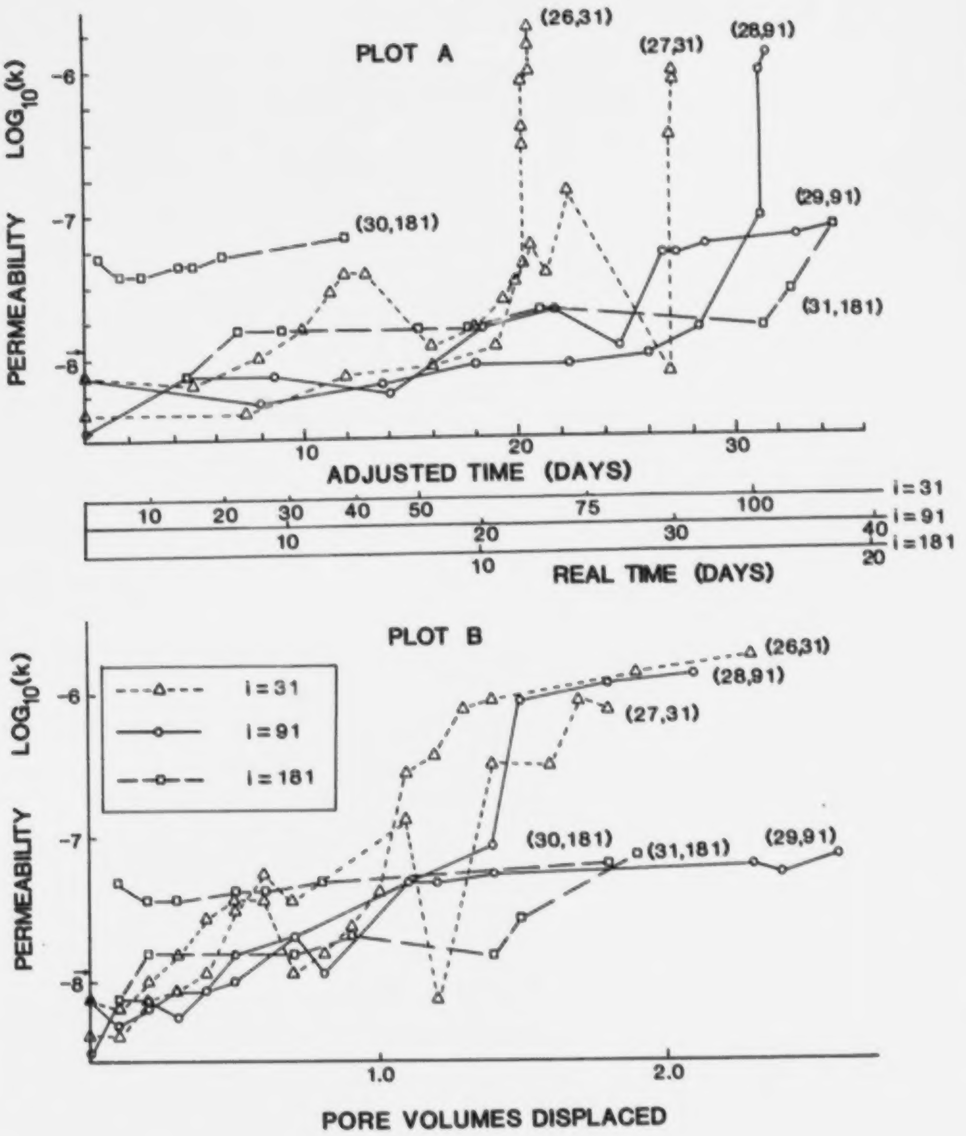
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figures 65-A and 65-B. Permeability of nonsaturated Elgin Kaolinite with xylene



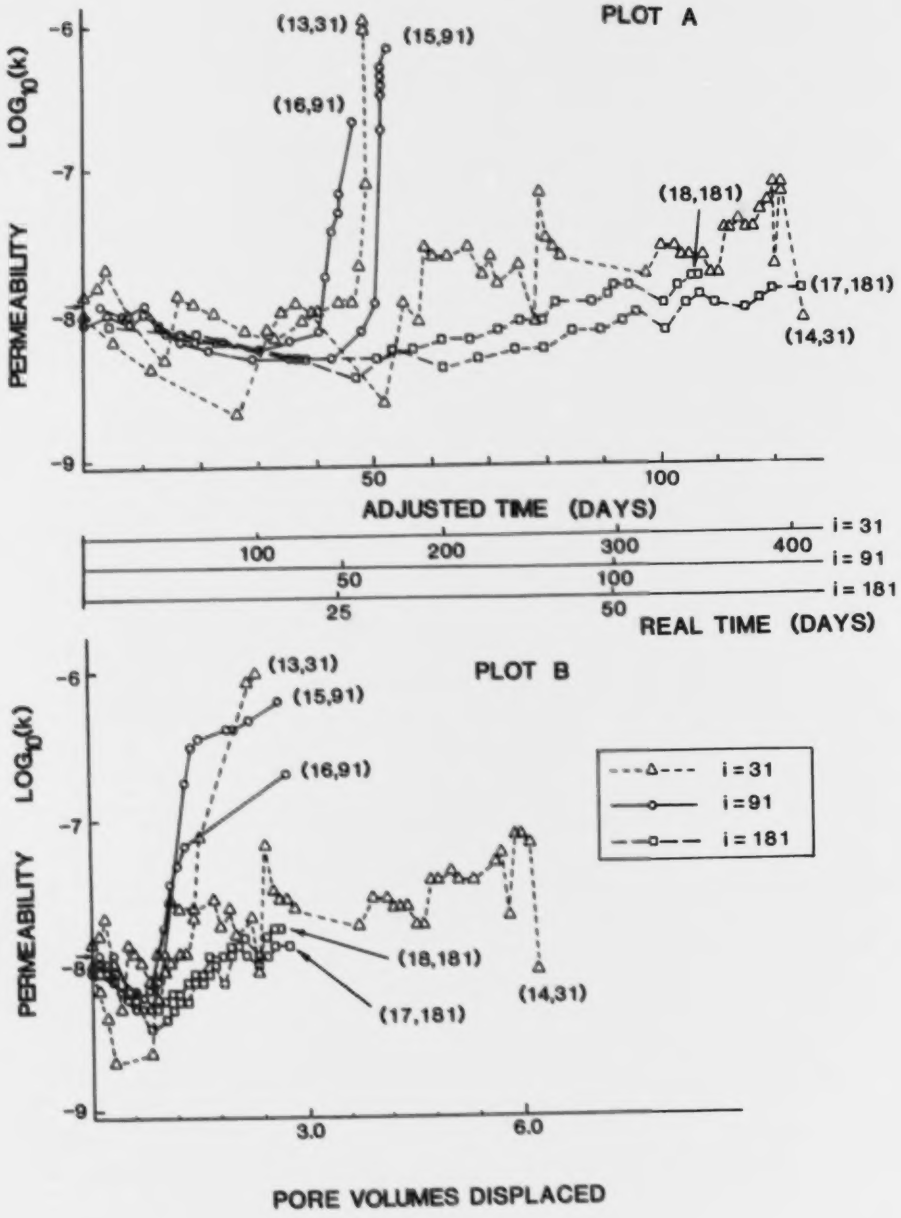
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figures 66-A and 66-B. Permeability of presaturated Elgin Kaolinite with xylene



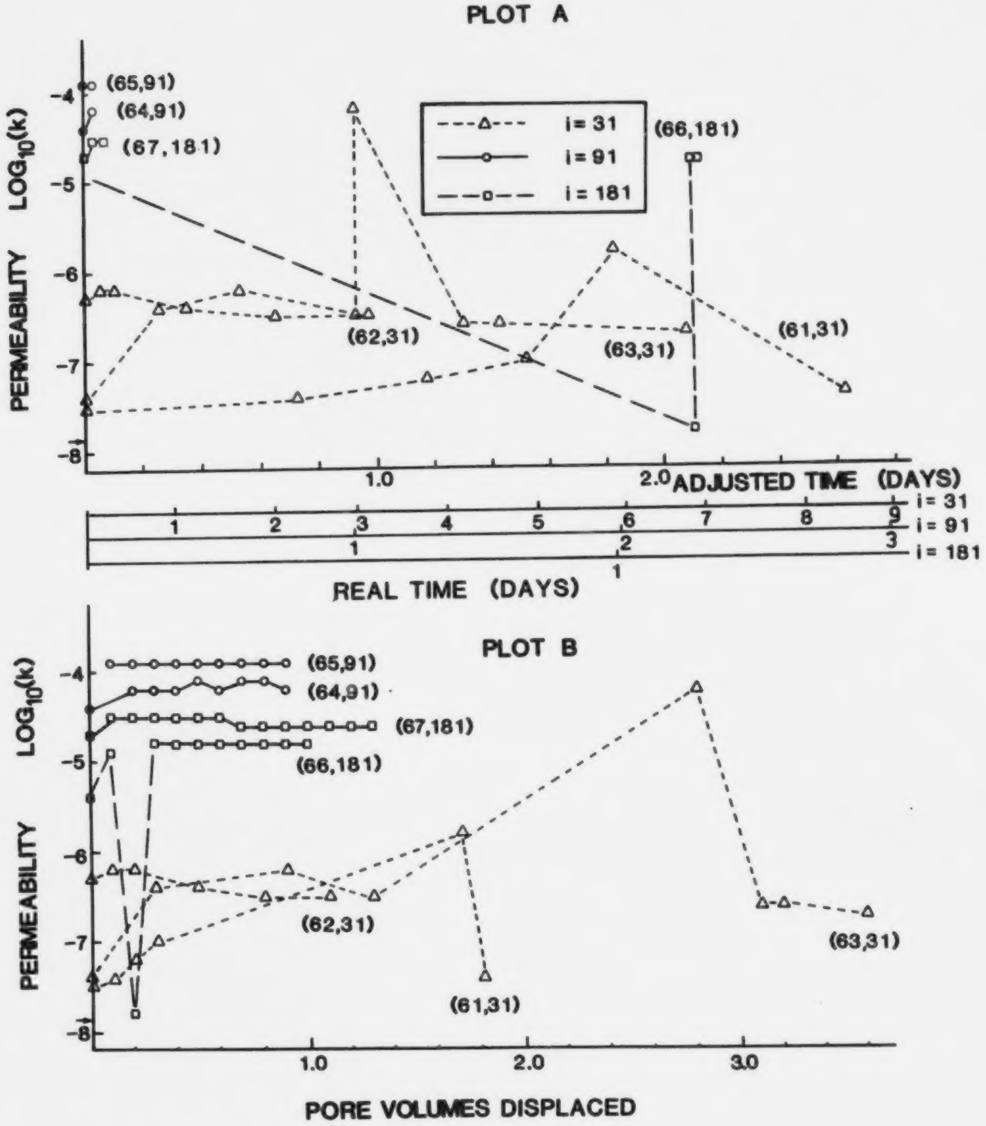
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figures 67-A and 67-B. Permeability of nonsaturated Micaceous Clay with acetone



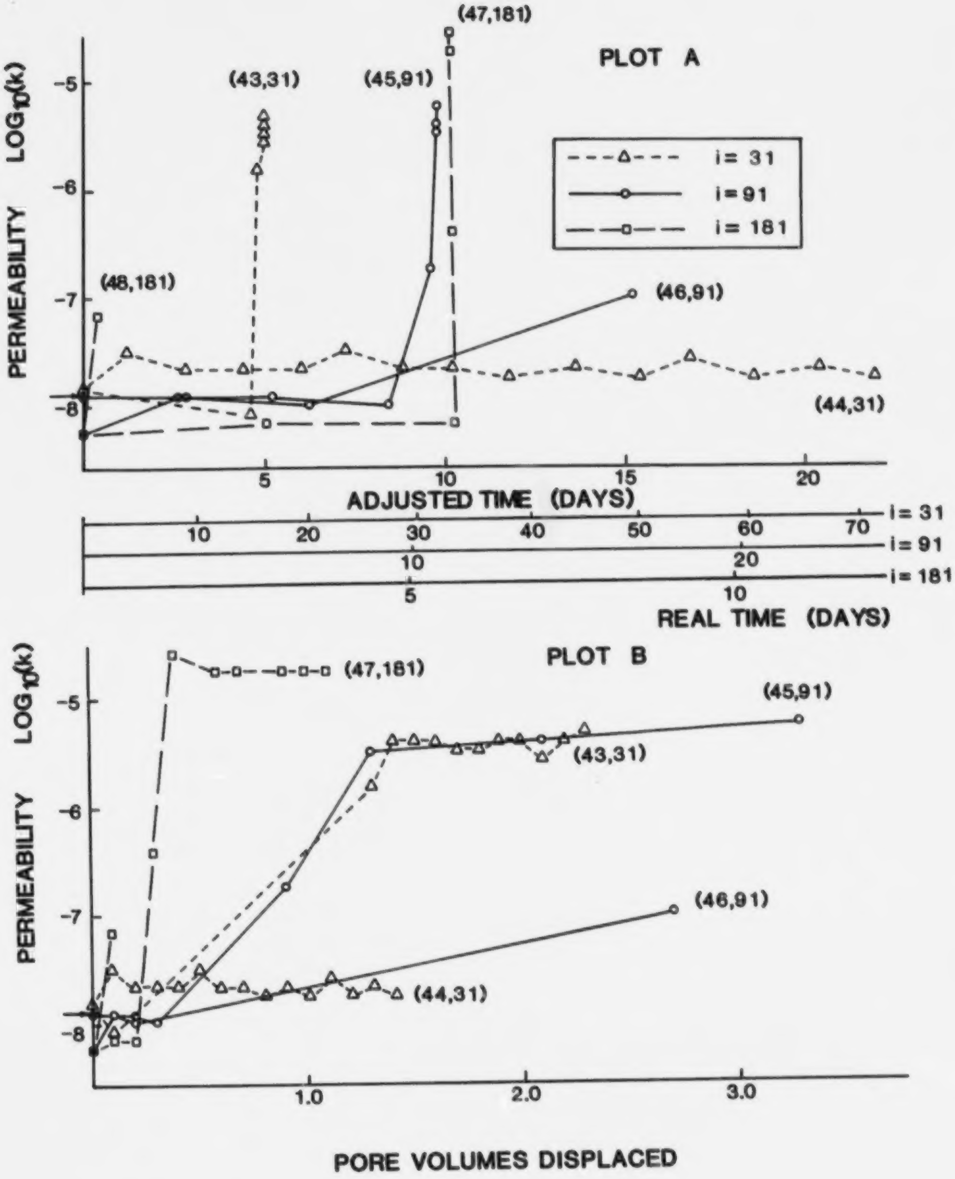
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figures 68-A and 68-B. Permeability of presaturated Micaceous Clay with acetone



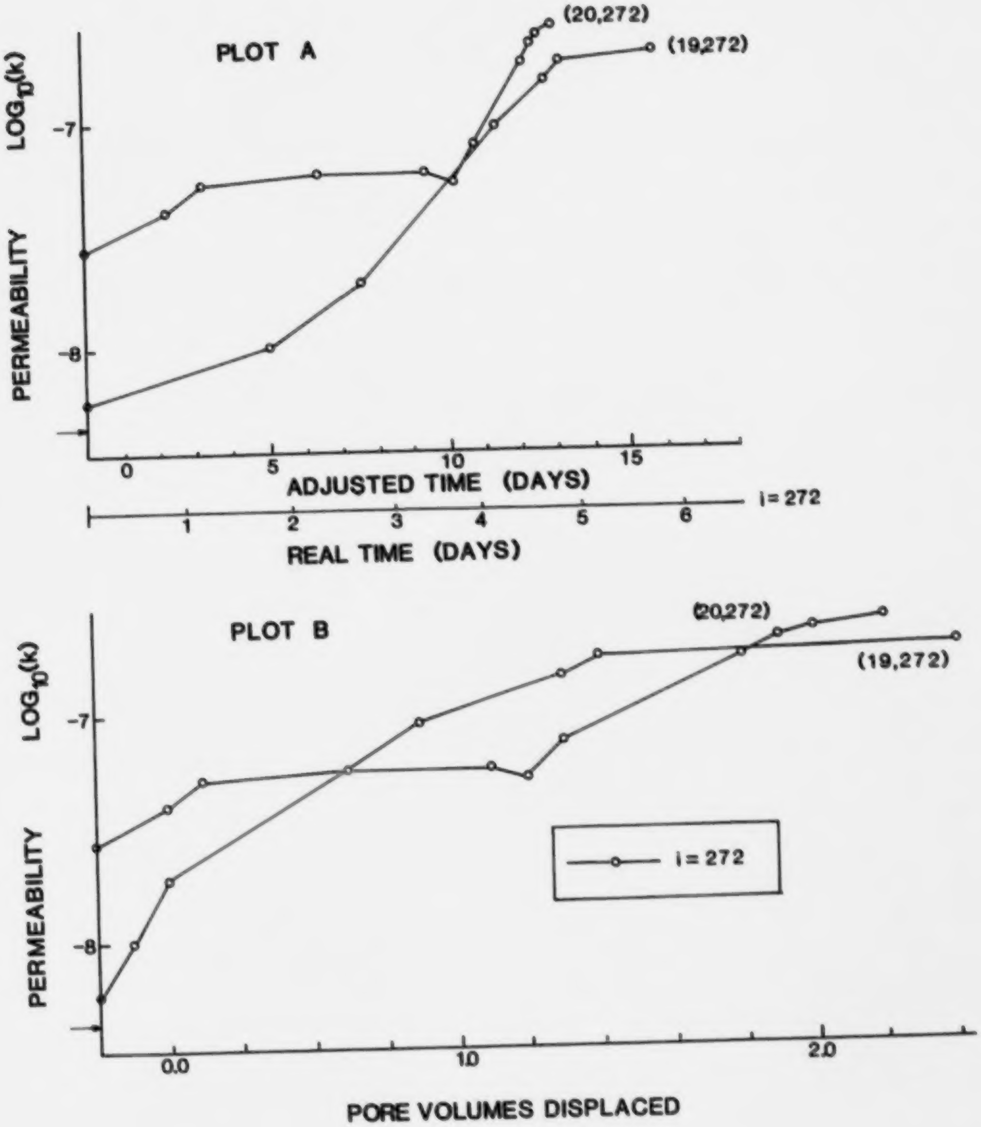
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figures 69-A and 69-B. Permeability of nonsaturated Micaceous Clay with xylene



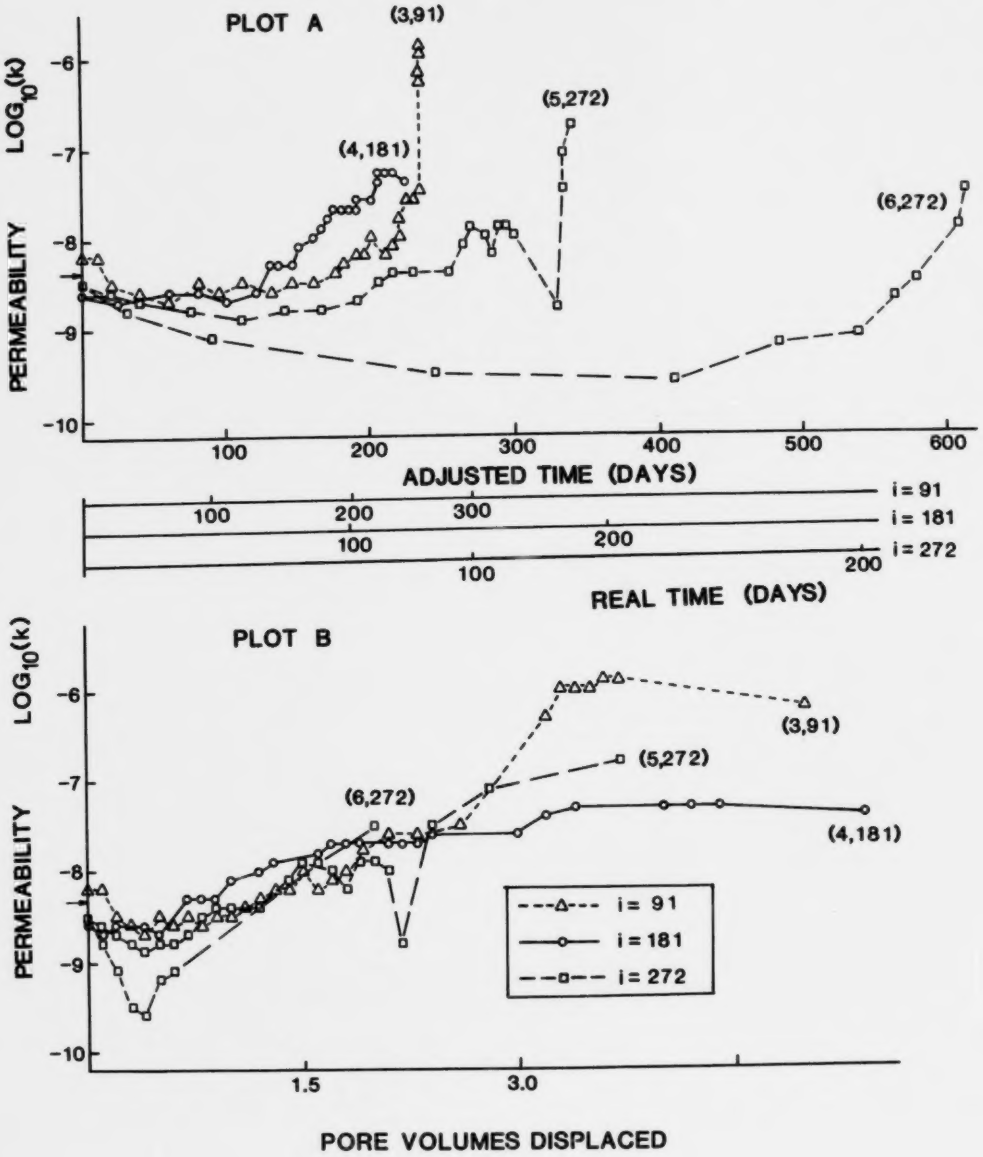
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figures 70-A and 70-B. Permeability of presaturated Micaceous Clay with xylene



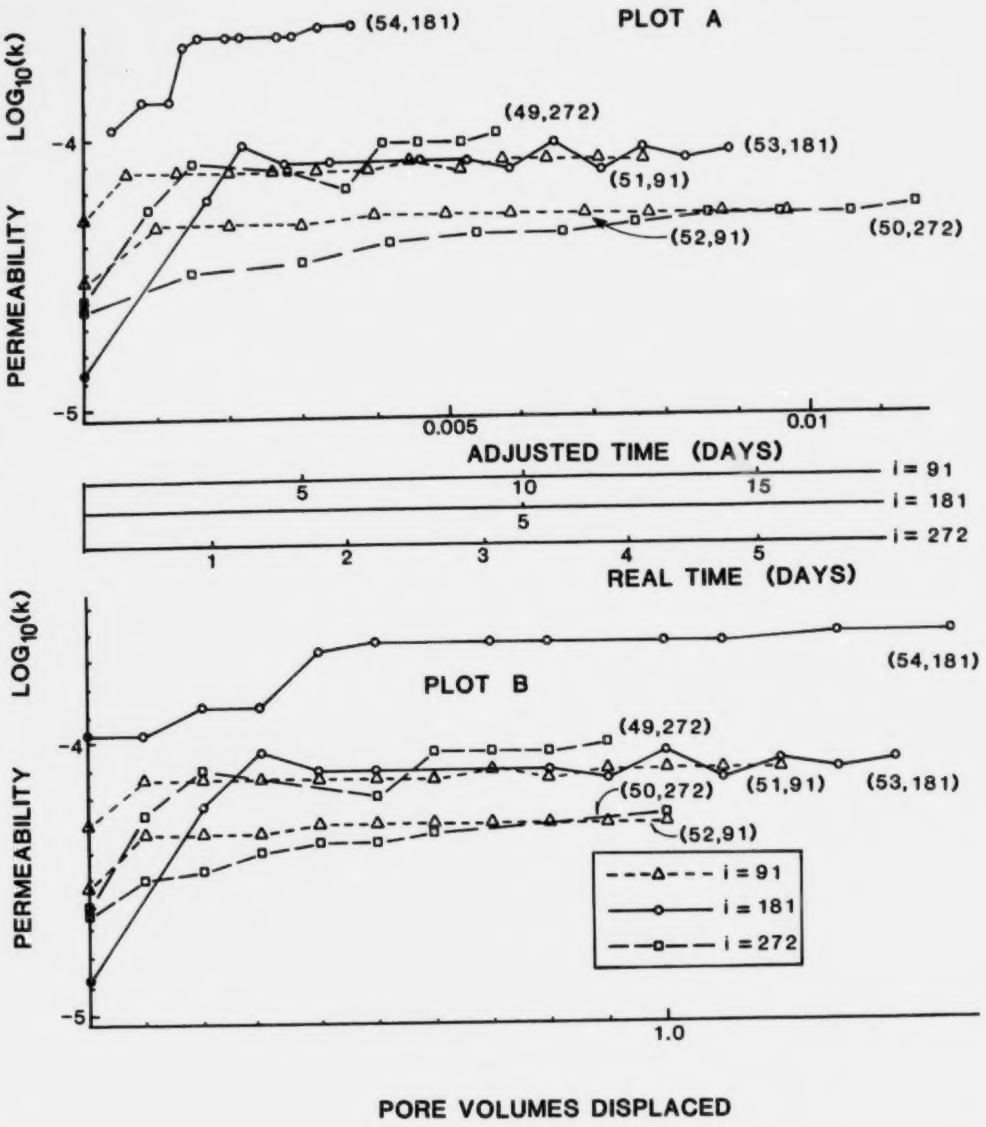
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figures 71-A and 71-B. Permeability of nonsaturated Bentonite Clay with acetone



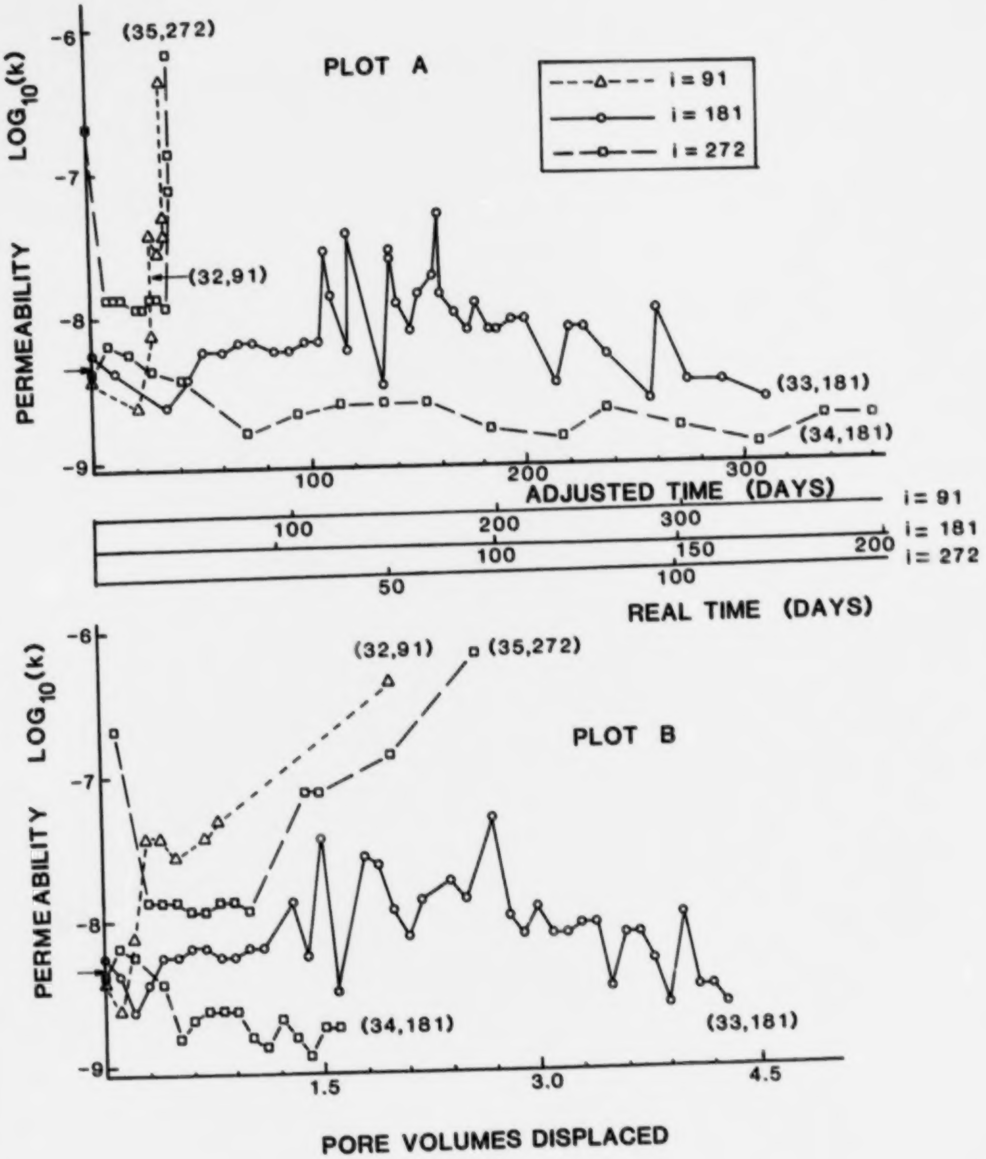
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figures 72-A and 72-B. Permeability of presaturated Bentonite Clay with acetone



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figures 73-A and 73-B. Permeability of nonsaturated Bentonite Clay with xylene



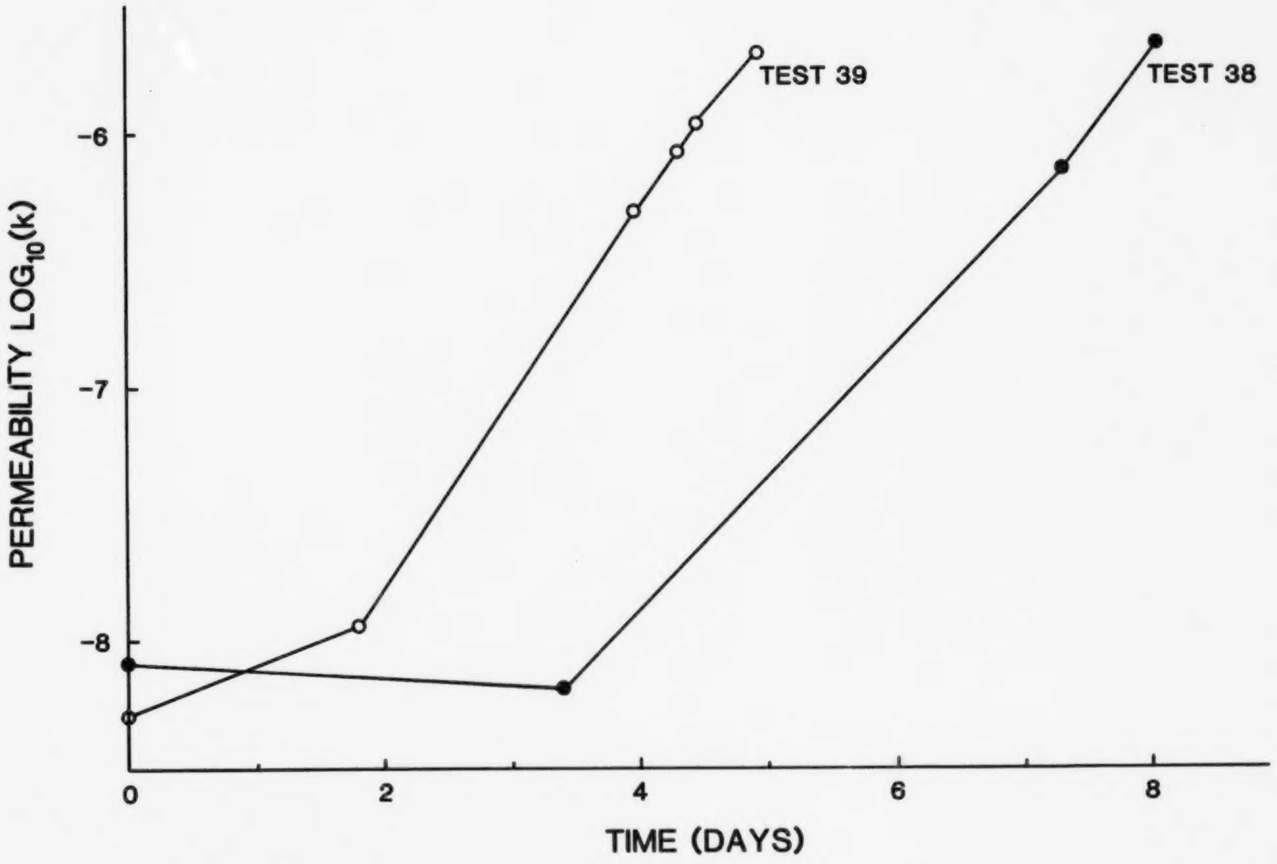
(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figures 74-A and 74-B. Permeability of presaturated Bentonite Clay with xylene test does appear flat over the last 2.6 pore volumes, but all this flow occurred over the last one and one-half days of the test, whereas the first 2.4 pore volumes of flow required more than 40 days.

In these writers' opinion, the behavior of these tests seems far more consistent when viewed in the perspective given by the time plot in Figure 64-A. It suggests that each of the specimens underwent a period of relative stability until the acetone chemically acted on the clay to create shrinkage cracks or channels. At that point, each of the specimens in essence failed, thereby permitting greatly increased flow. The fact that the terminal permeabilities recorded for the six specimens differed by a factor of two and one half orders of magnitude seems irrelevant from this point of view; these differences may to be more related to when the tests were discontinued than any other factor. Put another way, the apparent leveling off of tests such as these, when viewed in terms of pore volumes, may not relate to any physical state of equilibrium. For example, test 8 in Figure 64-B has remained stable for nearly one pore volume, but this corresponds to just 11 hours of test time. Considering that this test had just experienced an order of magnitude increase in permeability prior to this time interval, and knowing that these increases occur episodically, it seems difficult to conclude that this test has equilibrated. Subjectively, the writers have no difficulty imagining that eventually the permeability of this test might reach the levels shown in test 7, which itself has exhibited steady permeabilitiy over only the last 20 minutes of testing, during which time it passed one pore volume.

This discussion was not meant to suggest that rigid-wall tests do not equilibrate under any circumstances. Rather, it is suggested that in situations where concentrated chemicals of the nature used in these experiments are tested in the rigid-wall permeameter and are seen to cause the kinds of effects observed in Figure 64, it is at best very difficult and in some cases probably infeasible to attempt to run tests until they achieve steady conductivity. In the opinion of these writers this does not detract from the utility of the experiments, but does require the development of test measures for analysis which are not based on the notion of equilibrium, but rather reflect the idea of time to failure or of probability of failure within a given time interval, as defined according to specific and relevant failure criteria.

Figures 66, 68, 70, 72 and 74 summarize the behavior of the remaining presaturated tests in this data set, and are similar in many respects to Figure 64. They also illustrate the fact that time plots and pore volume plots can give very different impressions when tests are experiencing rapid permeability changes. Consider, for example, tests 38 and 39 in Figure 66 (Elgin Kaolinite Clay with xylene). The pore volume plot in Figure 66-B shows initial rapid increase in conductivity for these two tests, followed by an apparent tendency towards stabilization. The time series plot, on the other hand, shows no tendency to equilibrate on the part of either test and indicates that both have increased in permeability by two and one-half orders of magnitude within roughly the first week of testing. Time plots of these two tests on a more convenient scale are also shown in Figure 75.



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 75. Permeability of presaturated Elgin Kaolinite with xylene (Tests 38 and 39)

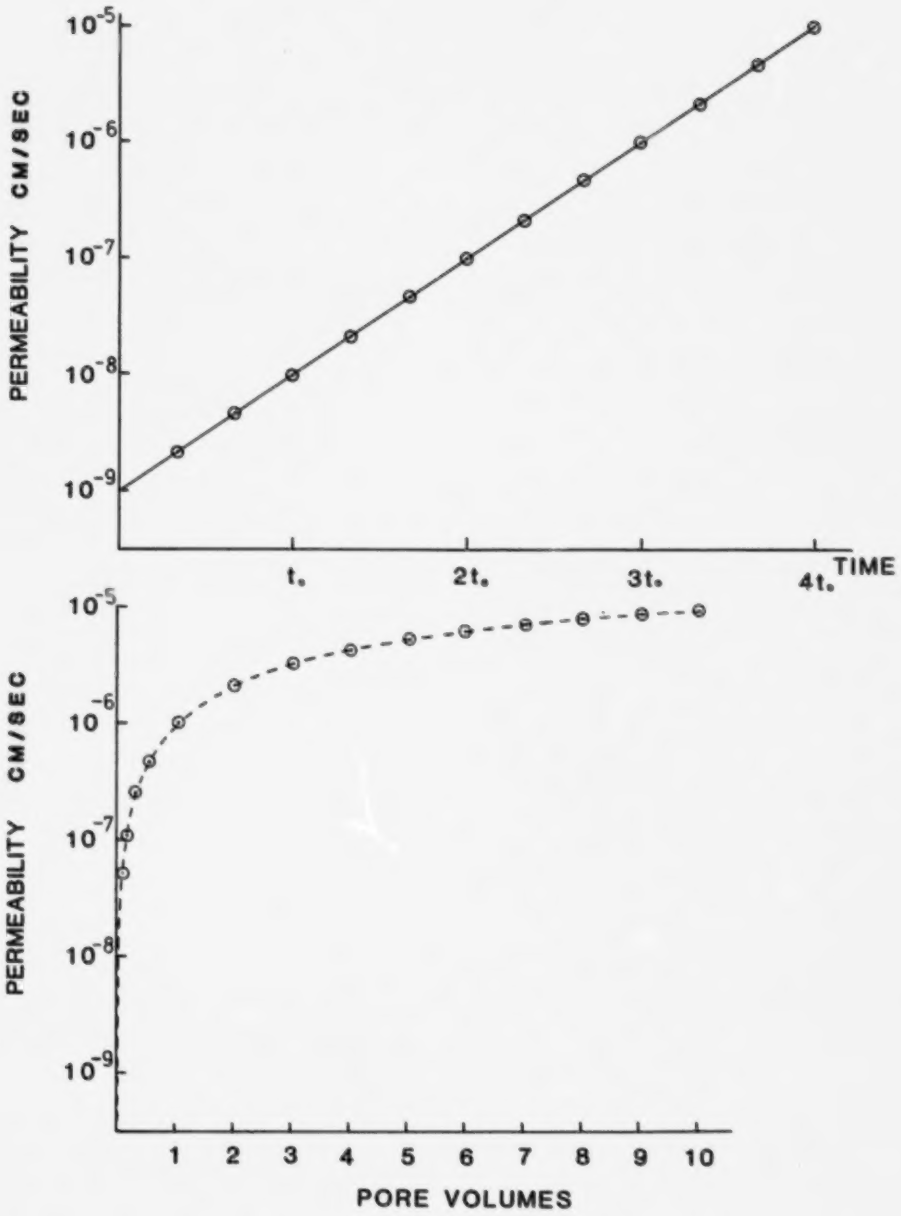
As a mathematical abstraction of this example, straightforward calculus shows that if over a given time interval permeability is increasing exponentially with time, then when viewed as a function of pore volumes of flow it will be linearly increasing. When graphed on semi-logarithmic paper, it follows that a time plot of permeability will be linearly increasing, whereas a pore volume plot of permeability will be logarithmic. The difference in appearance is shown in Figure 76, which depicts plots of a hypothetical test whose permeability is increasing at the rate of one order of magnitude each unit of time. These plots are similar to those of tests 38 and 39 in Figures 66-B and 75.

In situations where permeability is decreasing with time, exactly the opposite pattern occurs. Pore volume plots give the appearance of much more rapid decrease than do time plots.

Several examples of presaturated tests in Figures 64, 66, 68, 70, 72, and 74 indicate that there is considerable variation in the time required for presaturated specimens to reflect the chemical attack of the permeants on the soils, even taking gradient differences of the tests into account. In Figure 66-A, test 36 does not fail until the 172nd day, whereas test 37 experiences rapid increase in permeability after about 15 days, even though both were run at a gradient of 31. In the same way, Figure 72-A shows that test 6, in which acetone is permeated through Bentonite Clay, undergoes steadily decreasing conductivity over the first 150 days of testing, with rapid increases in permeability only after 200 days. The other three tests in this series experience meaningful increases in permeability in about half that time, after nominal adjustment for gradient differences.

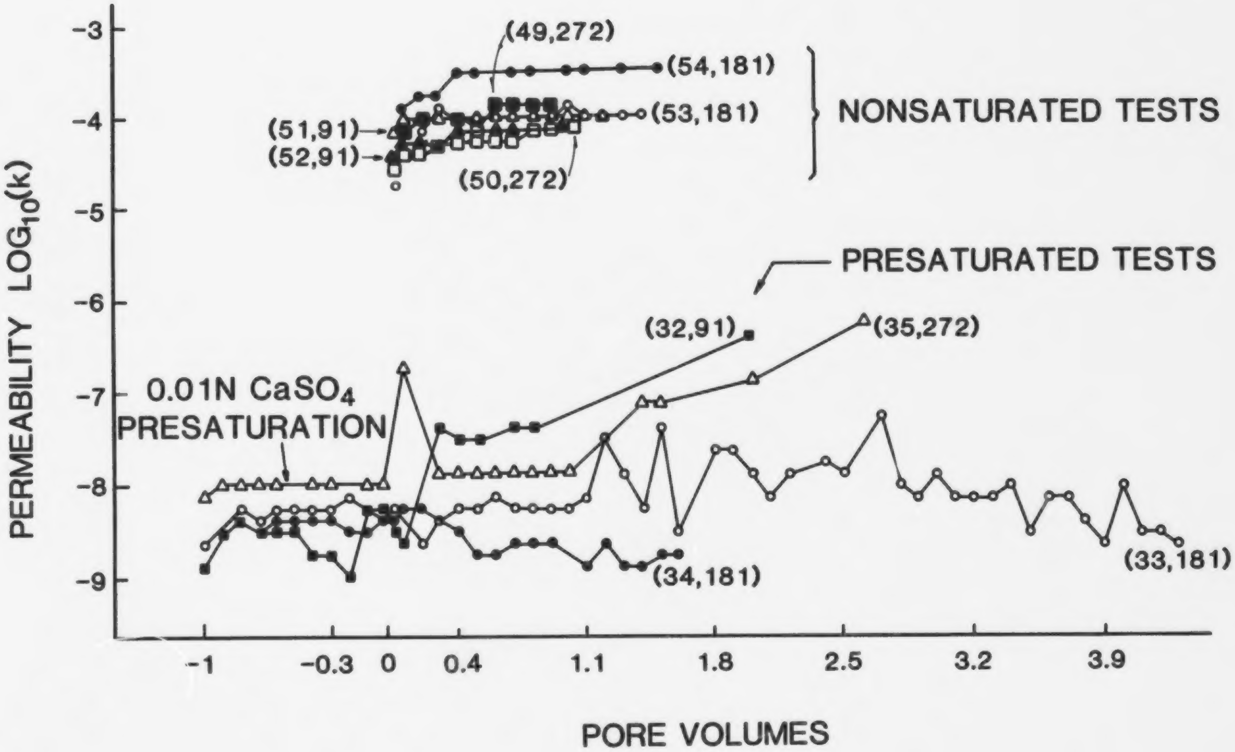
Although the general pattern displayed by the presaturated tests was one of eventual large increases in permeability, about one fifth of them did not display such behavior. The most graphic example of this is Figure 68-A, which illustrates tests of Micaceous Clay with acetone. Here, tests 13, 15 and 16, run at gradients of 31, 91 and 91 respectively, have each failed after about 45 days of time, adjusted to the benchmark gradient of 100. The remaining tests 14, 17 and 18, at gradients of 31, 181 and 181, have not. This does not seem particularly unusual if the experiments are thought of in analogy to product life tests (37) or (since tests are run at high gradients to substitute pressure for time) to accelerated life testing (37) since as a practical necessity such tests are commonly terminated before all units on test have failed. Moreover, the behavior of these tests may be due to spurious causes. After they were terminated, Brown noted that several of those specimens which had not experienced conductivity increases had deposited a residue which clogged the porous stone underneath the clay soil. This factor may well have caused artificially low permeabilities in these tests.

The differences between presaturated and nonsaturated tests were sometimes quite extreme. The most vivid example of this is shown in Figure 77, which compares the presaturated and nonsaturated tests on Bentonite Clay with xylene as permeant. The nonsaturated tests have initial permeabilities about four orders of magnitude above those of the presaturated tests, and as shown in Figures 73-A and 73-B, accumulate from .9 to 1.5 pore volumes in a matter of minutes (as previously noted, these times do not account for the



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure 76. Clay soil with exponentially increasing permeability



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG$_{10}$ PERMEABILITIES)

Figure 77. Permeability of presaturated and nonsaturated Bentonite Clay with xylene (Numbers in parentheses are test numbers and gradients.)

time between the start of the test and the initial permeability measurement). An inspection of Figures 63, 65, 67, 69 and 71 will show that, with the exception of Micaceous Clay permeated with acetone, it was typically the case that the nonsaturated tests run at higher gradients had initial permeabilities well in excess of those of the presaturated tests and accumulated the number of pore volumes required for test termination almost immediately. Nonsaturated tests run at lower gradients generally had lower initial permeabilities, which in some cases decreased to the water baseline level. If this occurred, the tests thereafter tended to behave somewhat like presaturated tests. Figure 63 illustrates these comments. Tests 24 and 25, which were run at a gradient of 181, initially had conductivity of $10^{-6.2} = 6 \times 10^{-7}$ cm/sec or more. The slower of the two accumulated 2.9 pore volumes in slightly more than one day. Tests run at gradients of 31 and 91 started at lower permeabilities and experienced decreases toward the general vicinity of the water baseline. Speculatively, this may have been caused by the fact that lower initial permeability coupled with less pressure have given these tests enough time to swell. After about 90 days test 21 broke through to the level achieved by the tests run at a gradient of 181. Tests 22 and 23 had not failed prior to termination.

SECTION 5

EXPERIMENTAL PLANS FOR PERMEABILITY STUDIES

5.1 INTRODUCTION

In this Section and in Appendices B and C, statistical aspects of the design of permeability studies are discussed. Section 5.2 describes several types of experimental plans which may be useful in permeability experiments whose goal is the comparison of equilibrium permeabilities obtained under application of various treatments, such as different chemical permeants, gradients or permeameter types. In Appendix B is found a brief introduction to the statistical analysis of data collected according to the plans described in Section 5.2. Appendix B also includes an outline of general considerations concerning the design of a testing program for the evaluation of potential borrow material to be used in liner construction. Appendix C discusses the problem of permeability test equilibration.

5.2 EXAMPLES AND COMPARISONS OF EXPERIMENTAL PLANS

In this Section a number of experimental plans or designs are discussed which may be of use in proposed laboratory research studies designed to compare the effect of various factors on permeability tests. It is not suggested that a single protocol is appropriate for all studies which might be proposed, since the best approach in a particular situation is specific to the research objectives of interest, and also is a function of the degree to which the experimenter is capable of controlling such extraneous factors as soil heterogeneity, molding water content and compactive effort. The discussion here is intended to describe the basic features of a number of potentially useful plans to compare their relative strengths and weaknesses, and to illustrate the concepts of randomization, replication and blocking.

Generally, the information contained in this Section is directed toward the non-statistician who is involved in the development of experimental plans. An experienced statistician should be consulted as required both during the experimental planning phase of the project and during the statistical analysis of the data resulting from the experiment.

5.2.1 References

Much of the material in this Section is discussed in detail in texts on the statistical aspects of experimental design. Given below is a short list of texts which may be of use in designing or evaluating a comparative permeability experiment:

Elementary, Applied Treatment:

Box, G.E.P., Hunter, W.G., and Hunter, J.S. Statistics For Experimenters. John Wiley & Sons, New York, 1978.

Hicks, C.R. Fundamental Concepts in the Design of Experiments. Holt, Rinehart and Winston, New York, 1973.

Intermediate, Applied Treatment:

Milliken, G.A., and Johnson, D.E. Analysis of Messy Data, Vol. 1: Designed Experiments. Lifetime Learning Publications, Belmont, CA, 1984.

Winer, B.J. Statistical Principles In Experimental Design, 2nd Ed. McGraw-Hill, New York, 1971.

Useful Compilation of Plans:

Cochran, W.G., and Cox, G.M. Experimental Designs, 2nd Ed. John Wiley & Sons, New York, 1957.

Theoretical Treatment:

Graybill, F.A. Theory and Application of the Linear Model. Duxbury Press, Belmont, CA, 1976.

Scheffe, H. The Analysis of Variance. John Wiley & Sons, New York, 1959.

Searle, S.R. Linear Models. John Wiley & Sons, New York, 1971.

5.2.2 Terminology

The terminology used here essentially follows that used in the text of Milliken and Johnson (34):

Treatment Structure: The treatment structure of an experimental design specifies the set of treatments or treatment combinations to be compared in the experiment. In the simplest case, the effects of a single factor at k different levels are to be compared. In this case the treatment structure is referred to as a "one-way" treatment structure. An example would be an experiment in which the permeabilities of a particular clay soil were compared at three levels of gradient, say 50, 100 and 150, by testing several samples at each of the indicated gradients while holding other relevant factors fixed. Thus the treatment structure consists of a one-way classification with three treatments.

In experimental designs in which two factors are investigated, one at "a" different levels and the other at "b" levels, a "two-way" treatment structure is formed by combining each level of the first factor with each level of the second factor. This results in (a)(b) treatment combinations to be investigated. Thus, an experiment simultaneously comparing the effect of gradient at levels 50, 100 and 150, using both rigid-wall and flexible-wall permeameters, has a two-way treatment structure where the first factor has three levels and the second factor has two levels, resulting in six treatment combinations.

Experiments having two-way, three-way or higher-way treatment structures are referred to as factorial experiments. Thus the previous example might be called a "3x2 factorial experiment."

Experimental Unit: In general, the experimental units of an experiment are the basic units of raw material to which treatments will be applied. In permeability studies, the experimental units will consist of quantities of clay soil obtained by dividing the available material into portions suitable for the application of treatments. The nature of these "quantities" and the method by which treatments are assigned to them depend strongly on the nature of the experimental design, as will be indicated below.

Replications: Replications of a treatment (or treatment combination) refer to independent trials in which the treatment is applied to different experimental units.

Design Structure: The design structure of an experimental design indicates how experimental units are to be grouped into "blocks" or homogeneous groups. Application of treatments to experimental units from within the same block then permits comparisons under conditions which are reasonably uniform.

In situations where experimental units are quite homogeneous to begin with, blocking of experimental units may not be required. In this case treatments should be assigned to experimental units completely at random, subject to the constraint that each treatment is replicated the desired number of times. This design structure is referred to as a "completely randomized design".

Consider an experiment whose treatment structure calls for k treatments (or treatment combinations). The simplest design structure involving blocking requires the grouping of experimental units into homogeneous blocks of k experimental units apiece. If the k treatments are then assigned at random to the k experimental units within each block, the resulting design structure is called a "randomized complete block design".

In the following Sections the set-up of completely randomized and randomized complete block designs will be discussed and illustrated. In addition, randomized incomplete block and split-plot designs will be briefly considered. Randomized incomplete block designs are sometimes useful in situations where the number of treatments to be compared is too large to accommodate all treatments within each block of a randomized complete block design without use of block sizes which are too large to maintain adequate within-block homogeneity of experimental units. Thus designs are considered in which not all treatments are applied to experimental units within every block. Split-plot designs are a special case of incomplete block designs which have (at least) a two-way treatment structure, and which employ different sized experimental units to which the treatments are applied. For reasons which will be discussed below, certain comparative permeability experiments require such a design.

The process of experimental design requires the determination of the following aspects:

1. the factor(s) to be investigated
2. the treatment structure to be utilized
3. the nature of the experimental units
4. the design structure used to block these experimental units
5. the method by which the assignment of treatments to experimental units is to be randomized, consistent with the treatment and design structures
6. the amount of replication required to achieve adequate comparison of treatments.

These points should be specified in experimental plans submitted to EPA. Since an understanding of points 1 through 5 is also imperative for the proper statistical analysis of experimental data, they should also be documented in final reports. Thus it is strongly recommended that proposed research projects be designed and analyzed in collaboration with an experienced experimental statistician.

5.2.3 Completely Randomized Design

Consider an experiment in which it is desired to compare the effects of $k=4$ different chemicals C_1, C_2, C_3 and C_4 on the permeability of a particular clay soil. For purposes of the example, suppose that other relevant factors are held fixed (that is, the same type of permeameter is to be used throughout, all test specimens will be moisterized to the same nominal molding water content, and so forth) and that $n=6$ replications of each chemical are required. Thus the treatment structure is a one-way classification with four treatments.

Suppose it is decided that the experiment should be conducted by use of a completely randomized design structure. In this case, the clay soil available for the experiment should be divided into twenty-four batches, each of sufficient size to accommodate one test. These batches then constitute the experimental units for the experiment. The four chemicals are then randomly assigned to batches, in such a way that all possible assignments are equally likely, subject to the constraint that each chemical be assigned to six batches. This can be accomplished by use of a random number table, a random number generator, or specialized software such as SAS's PROC PLAN (27). The batches should be separately moisterized to the specified nominal molding water content, since replication is to be accomplished on independent experimental units. The order in which the 24 batches are processed (moisterized, compacted into cells and placed on test) should also be randomized.

The purpose of random assignment of treatments to experimental units is two-fold. First, it prevents between-batch differences in soil characteristics from systematically biasing the comparison of chemical effects, and in addition it tends to make the usual statistical analysis procedures less sensitive to violations of the mathematical assumptions on which they are formally based. Similarly, the randomization of the order of processing prevents influences which may be time-dependent (such as improvement in moisterizing or compaction techniques) from systematically entering into treatment comparisons.

A brief outline of analysis of variance procedures for completely randomized designs is found in Appendix B.2.

Use of a completely randomized design structure in this example causes differences in the soil characteristics of the batches to reduce the precision with which treatments may be compared. Similarly, deviations from the nominal molding water content of the batches, and differences in the efforts applied in their compaction, all contribute to random error. Thus use of a completely randomized design structure is not advisable unless the soil available for experimentation is quite homogeneous, or can be made so by mixing prior to initiating the experiment. In the same way, the design is probably not adequate unless the batch-to-batch variation of molding water content can be well-controlled, particularly if tests are to be conducted at a nominal molding water content close to optimum (for density) where the effect of even small changes in water content on permeability can be quite pronounced. If these conditions are not met, the experimenter should consider a design structure in which experimental units are blocked, to improve the uniformity under which treatments are compared. Such a design structure is discussed in the next section.

5.2.4 Randomized Complete Block Design

By use of a randomized complete block design structure, the experiment described in Section 5.2.3 may be altered so as to permit the comparison of chemicals under more uniform conditions. To do this, the clay soil available for the experiment is partitioned into six batches, each of sufficient size to accommodate four tests. Each batch is in turn moisturized to the same nominal molding water content and then quartered, resulting in four sub-batches, each of which is then randomly assigned to one of the four chemicals. Within each batch, the order in which the tests are set up is randomized. Thus in this experiment the treatment structure is a one-way classification with four treatments as before, but now the experimental units are sub-batches which are grouped within six blocks (commonly moisturized batches).

The rationale for blocking in commonly moisturized batches is based on three points:

1) Soil characteristics should be more uniform within batches than between batches.

2) Molding water content should be more uniform within batches than between batches.

3) Tests compacted at nearly the same time should be more uniformly prepared than tests compacted at different times.

The rationale for within-block randomization of the assignment of sub-batches to chemicals, and for the order of their preparation, is exactly as before. Extraneous factors not accounted for by between-block differences are prevented from systematically biasing treatment comparisons, and the analysis is made less sensitive to violations of formal assumptions.

Appendix B.3 contains a brief outline of analysis of variance procedures to be applied to data obtained from experiments which employ randomized complete block designs.

A comparison of the randomized complete block design and the completely randomized design shows that the randomized complete block design controls for unavoidable between-batch differences. Thus, unless between-batch differences in soil characteristics, molding water content and specimen preparation are negligible, such blocking permits sharper comparisons of treatments.

The elimination of batch differences entails the loss of n-1 degrees of freedom for error (Appendix B.3). If degrees of freedom are plentiful, the effect of this loss is small. In such cases, the replication of experiments in randomized complete blocks is highly recommended for comparative experiments when feasible. In situations where the degrees of freedom are required to be small for reasons of economy, the loss of degrees of freedom to eliminate block effects could offset the reduction in experimental error obtained by blocking, if between-block differences are small. The decision whether to use a completely randomized or randomized complete block design structure then should be based on the evaluation of the amount of replication required by both designs to achieve adequate power in treatment comparisons, as discussed in Section 5.2.6. and Appendix B.6.

The general concept of blocking has been discussed here in terms of comparison of treatments within commonly moisturized batches of clay soil. Blocking may be useful in other contexts as well. For example, whenever feasible, comparative tests should be run concurrently to eliminate possible environmental effects. In experiments having a large number of treatment combinations, it may not be possible to do this, due to limitations on available equipment or space. In this case, the experiment may be replicated in randomized complete blocks of concurrently run tests.

5.2.5 Factorial Experiments

Experiments having a two-way or higher-way treatment structure may be replicated by using either a completely randomized or randomized complete block design structure. The concepts extend in a straightforward way from those discussed in Sections 5.2.3 and 5.2.4.

Factorial experimentation is of interest because it is generally more efficient than separate investigations of isolated factors, and because it lends itself to the estimation of possible interactions in the effects of factors considered in combination. Two-way or higher-way treatment structures may also be recommended when the principal objective of a study lies in the determination of the effects of one specific factor, but for purposes of promoting the generality of results it is deemed advisable to investigate this factor under a variety of conditions. For example, a study may be proposed whose purpose is to explore possible differences in measured permeability of a clay soil obtained by use of rigid-wall and flexible-wall permeameters, but for the sake of generality it is decided to conduct tests using each of three different chemical permeants.

The example described in the previous paragraph leads to a two-way treatment structure, having two factors, permeameter type (P) and chemical (C) at $p = 2$ and $c = 3$ levels respectively. Suppose that the experiment is planned to include $n = 5$ replications of each of the $(pc) = 6$ treatment combinations. If this experiment is to be replicated by use of a completely randomized design structure, then the clay soil available for experimentation would be divided into 30 batches each of size sufficient to accommodate one test. The six treatment combinations (permeameter-chemical combinations) would then be randomly assigned to these batches by use of a random number table, random number generator or appropriate software, in such a way that each treatment combination is assigned to five batches. Each batch is then independently moisturized to the specified nominal molding water content and the corresponding test prepared and set up, with the order in which the 30 batches are processed randomized.

In this example, each treatment combination is replicated the same number of times. While this is not necessary, in Appendix B.7 some reasons are given which indicate that such balanced replication is generally to be recommended.

An alternative method of performing this factorial experiment would involve replication of the six treatment combinations in complete blocks. To do this, the clay soil available for the experiment would be divided into five batches of sufficient quantity to accommodate six tests apiece. These batches are in turn moisturized and then divided into six sub-batches which are randomly assigned to one of the six permeameter-chemical combinations. Within each batch, the order in which the six tests are set up is randomized.

The motivation for replication in complete blocks is that between-batch differences in soil characteristics, molding water content and so forth are eliminated from treatment comparisons, thus permitting sharper tests of their effects. Thus, just as in Section 5.2.4, such blocking is recommended unless the material available for testing is extremely homogeneous, and strict control of molding water content and compactive effort is possible from test to test.

Analysis of variance procedures for factorial experiments are described in Appendices B.4 and B.5.

5.2.6 Determining the Required Number of Replicates

In planning an experiment, it is necessary to decide on the number of times each treatment must be replicated in order to provide useful experimental results. The most common approach to this problem is to specify one or more null hypotheses concerning the comparison of treatments and to provide sufficient replication such that these tests will have a prescribed power (that is, probability that the null hypothesis will be rejected) against pre-chosen alternative hypotheses.

A second method of sample size determination is available if the experimenter has apriori a particular interest in certain specific treatment comparisons. If so, the number of replications may be determined such that the widths of confidence intervals for these comparisons do not exceed specified bounds. The specific details of the two approaches described above are outlined in Appendix B.6, and numerical examples are also given there.

5.2.7 Incomplete Block Designs Which Confound Factorial Effects

The replication of experiments in randomized complete blocks requires that each block contain as many experimental units as there are treatment combinations to be investigated. In some cases this may be undesirable, because it may lead to block sizes which are too large to maintain reasonable uniformity of experimental units.

As an example, consider a factorial experiment whose purpose is to compare the permeabilities of a particular clay soil using two different permeants (e.g. standard aquaeous leachate and methanol) and two permeameter types (rigid-wall and flexible-wall) at each of four different gradients (50, 100, 150, 200). Suppose that each treatment combination is to be replicated three times, and that the experiment is to be replicated in randomized complete blocks consisting of commonly moisterized batches of the clay soil, which are subdivided into sixteen sub-batches to be randomly assigned to each treatment combination.

The experiment described in the previous paragraph may prove to be unsatisfactory, because the size of the batches required is so large. The moisterization of a batch large enough to accommodate sixteen tests may be too unwieldy to be practical, or it may not be feasible to prepare such a batch without introducing significant variation in soil characteristics and water content among the sixteen sub-batches. If this is the case, the experimenter may wish to consider a design structure which does not require such a large block size. One such approach is to make use of an incomplete block design, as discussed in this Section. Another possible approach is discussed in Section 5.2.8.

In an incomplete block design, not all treatment combinations are applied to the experimental units within any one block. This permits the experimenter to make use of smaller blocks, whose experimental units may therefore be kept homogeneous within each block. The cost of doing this is that, in any replication of the experiment, the comparison of certain treatment combinations is confounded with (that is, cannot be differentiated from) between-block differences. For this reason, if the use of an incomplete block design structure is under consideration, the experimenter must carefully consider its formulation to ensure that the information which is lost through confounding will not lead to experimental results whose interpretation is ambiguous.

A complete description of the formulation and analysis of factorial experiments with confounding is beyond the scope of this report. The intention here is to provide an introduction which outlines the general features of such designs by means of an example. For a more detailed treatment of the subject, the reader is referred to (5), (48) and (10). The latter reference also includes experimental plans which are useful in choosing a design which is appropriate in a given situation.

Consider again the factorial experiment investigating the three factors permeant (P, at two levels), permeameter type (T, at two levels) and gradient (G, at four levels) which is to be replicated three times. Suppose the experimenter determines that while the use of batches large enough to accommodate sixteen tests apiece is not feasible, it would be practical to prepare commonly moisterized batches which are of sufficient size to accommodate eight tests, while still maintaining acceptable uniformity of soil characteristics and molding water content within each batch.

Each of the three replications of the experiment consists of sixteen tests, one for each of the possible treatment combinations. These must be compared within two batches, each including eight of the treatment combinations. Therefore, within each replication, differences between the two batches will be confounded with the comparison of the two sets of eight treatment combinations tested within each batch. If forethought is given to the allocation of treatment combinations to the batches, this may be accomplished in such a way that main effects and two-factor interactions are not confounded with batch differences. The three-factor interaction between P, T, and G may be decomposed into (4-1)(2-1)(2-1) = 3 components, each corresponding to one degree of freedom. Within each replication of the experiment, the sixteen treatment combinations may be assigned to the two batches in such a way that one of these three single degree of freedom components is completely confounded with batch differences, while the remaining two are not confounded. It is generally good practice to choose different components to be confounded in subsequent replications of the experiment. This strategy, which is referred to as partial confounding, permits within-batch information to be obtained on each component within different replications of the experiment.

Suppose the two levels of the factor P are labeled 0 and 1, the two levels of the factor T are labeled 0 and 1, and the four levels of the factor G are labeled 0, 1, 2, 3. Then, the sixteen possible treatment combinations may be denoted by (0,0,0), (0,0,1), (0,0,2), (0,0,3), (0,1,0), (0,1,1), ..., (1,1,3), where the first digit refers to the level of factor P, the second to the level of factor T, and the third to the level of factor G. Thus, for example, (0,1,2) refers to a test conducted at level 0 of P (standard aquaeous leachate), level 1 of T (flexible-wall permeameter) and level 2 of G (gradient of 150). Now suppose that the treatment combinations are assigned to batches according to the plan outlined in Table 4:

The plan in Table 4 has the following features:

a. Each of the sixteen treatment combinations appears exactly once in each replication of the experiment.

b. Each level of any of the three factors appears the same number of times within each batch. This ensures that no main effect is confounded with between-batch differences.

c. Each possible combination of the levels of any two of the three factors appears the same number of times within each batch. Thus, no two-factor interactions are confounded with between-batch differences.

Table 4

INCOMPLETE BLOCK DESIGN
AN EXAMPLE

Replication #1		Replication #2		Replication #3	
Batch #1	Batch #2	Batch #3	Batch #4	Batch #5	Batch #6
(000)	(001)	(001)	(000)	(000)	(002)
(003)	(002)	(003)	(002)	(001)	(003)
(011)	(010)	(010)	(011)	(012)	(010)
(012)	(013)	(012)	(013)	(013)	(011)
(101)	(100)	(100)	(101)	(102)	(100)
(102)	(103)	(102)	(103)	(103)	(101)
(110)	(111)	(111)	(110)	(110)	(112)
(113)	(112)	(113)	(112)	(111)	(113)

d. The three degrees of freedom for the three-factor interaction between permeants, permeameter types and gradients are partially confounded. One of the three degrees of freedom is confounded in each of the three replications of the experiment, as described in Appendix B.8. Thus information on each component of the three-factor interaction is available from the two replications of the experiment in which it is not confounded with batch differences.

Construction of an analysis of variance table for this experiment is outlined in Appendix B.8.

As usual, the assignment of sub-batches to treatments, and the order in which tests are set up, should be randomized within each of the six batches shown in Table 4. In addition, the order in which the two batches are prepared within each replication of the experiment should be randomized.

As a general rule, the use of confounded designs in permeability research investigations should be approached with considerable caution. There are often good reasons to believe that the effects of certain factors may involve substantial interaction (e.g., the effect of increasing gradient in rigid-wall permeameters might very well differ from the effect of increasing gradient in flexible-wall permeameters under some circumstances). For this reason it is recommended that, if an incomplete block design is used, no interaction should be allowed to be completely confounded. Rather, partial confounding schemes should be used to obtain some within-batch information on all interactions.

Furthermore, the analysis outlined in Appendix B.8 assumes that the effect of batches is additive. While this assumption is tenable if batch preparation is reasonably uniform, the interpretation of experimental results is not straightforward if the assumption is violated. Therefore, the use of confounded designs should be considered only if the experimenter is convinced that use of a simpler design would not be suitable to achieve his or her purposes, and that the reduced block size obtainable by incomplete blocking will result in a meaningful reduction in within-block variability of experimental units. If use of the simpler designs described in Sections 5.2.3 through 5.2.5 is not appropriate, use of split-plot designs, to be described in the next Section, in some circumstances might still be preferable.

5.2.8 Split-Plot Designs

Consider an experiment whose purpose is the investigation of the effect of gradient on the permeability of a particular clay soil. Tests are to be conducted at gradients of 50,100 and 150, and to increase the generality of the experimental results it is decided to also vary the nominal molding water content at which the clay soil is compacted. Two levels of molding water content are chosen for experimentation, 15% and 20%. It is determined that each of the 3 x 2 = 6 treatment combinations are to be replicated five times.

Although this 3 x 2 factorial experiment could be replicated according to a completely randomized design structure as in Section 5.2.3, such an approach would lead to relatively imprecise comparisons of the effects of the three levels of gradient, due to between-batch variation in soil characteristics and actual molding water content. On the other hand, replication of the experiment in randomized complete blocks of commonly moisterized soil, as in Section 5.2.4, is not physically possible, by the very nature of the second factor of the experiment. Such a situation is not uncommon in permeability research experiments. For example, a similar problem would occur in an experiment at which the permeabilities of two different types of clay soil (e.g., a Kaolinite and a Bentonite) were to be determined at each of three different gradients.

A method of obtaining within-batch comparisons of gradients in this experiment is as follows: The clay soil available for the experiment is divided into ten batches, each of size sufficient to accommodate three tests. The two levels of nominal molding water content are then randomly assigned to these batches, such that each level is assigned to five of the batches. The batches are then moisterized, in a randomized order, to the indicated nominal water content, and divided into three sub-batches. For each batch, the three gradients are randomly assigned to sub-batches. Within each batch, the three tests are set up in random order.

This experiment has a two-way treatment structure, but actually has two distinct design structures. With respect to the molding water content factor, the design structure is completely randomized, since levels of this factor are assigned at random to the ten batches. Note that with regard to the water content factor, the experimental units are batches. On the other hand, with regard to the gradient factor, batches serve as blocks, sub-batches are the experimental units, and the design structure is a randomized complete block arrangement, since each level of gradient is assigned at random to a sub-batch within each of the ten batches used in the experiment.

An experiment of this type is referred to as a split-plot design. In the present context it would be more descriptive to call it a "split-batch" design, since whole batches, which serve as the experimental units for the water content factor, are split into sub-batches to provide the experimental units for the gradient factor. The term split-plot derives from the extensive use of such designs in agricultural experiments, in which plots of land are used as experimental units. In this context, the levels of one factor (e.g. different fertilizers) are applied to whole plots, which are further sub-divided into sub-plots to provide the experimental units for comparison of the levels of a second factor (e.g. different varieties of a particular crop).

Many authors would refer to designs of this nature as "repeated measures" designs, a term which derives from experimentation in the medical and behavioral sciences in which subjects are used as experimental units. In such cases subjects are often randomly selected from certain groups (e.g., categorical assessment of the severity of a certain medical condition) thus essentially forming a "between subject" factor, and each subject is assigned all levels of a second "within-subject" factor (e.g., different drugs to be compared with respect to their efficacy in providing symptomatic relief). Here, subjects are analogous to plots, individual observations on subjects are analogous to sub-plots, the between-subject factor corresponds to the whole plot factor, and the within-subject factor corresponds to the sub-plot factor. In the context of permeability testing, the present authors prefer the term "split-plot design" since in such testing batches are in fact physically split into sub-batches to form the experimental units for the sub-batch factor. Also, the term "repeated measures" is restricted by some authors (e.g. (34)) to experiments in which there are physical limitations on the ability to randomize the order in which within-subject treatments are applied, which is not the case here.

Returning to the illustrative example, there are two important points which distinguish this split-plot design:

a. There are two sizes of experimental units: For purposes of comparing the effects of molding water content (the whole plot factor) the experimental units are batches (whole plots). For purposes of comparing gradient effects (the sub-plot factor) the experimental units are sub-batches (sub-plots) within batches. Thus two different error estimates are required in testing for the effects of these factors, one which is a measure of batch-to-batch variability, and another with is a measure of within-batch variability between sub-batches. Failure to recognize this (e.g. treating the data as though the experiment had been completely randomized) will lead to an incorrect analysis which overestimates the significance of differences due to water content and underestimates the significance of gradient differences.

b. Two separate levels of randomization are required in preparing the experiment. In this example, batches are assigned levels of water content according to a completely randomized design structure, and gradient levels are assigned to sub-batches in accordance with a randomized complete block design structure.

An analysis of variance table for the split-plot experiment described above is given in Appendix B.9.

In the example presented in this Section, the use of a split-plot design was motivated by the fact that within-batch comparisons of different levels of molding water content is not physically possible. However, split-plot designs may be very useful even in cases where this is not so. In such cases, the experimenter has flexibility in determining which factor(s) will be compared within batches (sub-plot factors) and which will be compared between batches (whole plot factors). As a general rule, the experiment should be designed so that the factor(s) of primary experimental interest are compared within batches, thereby resulting in more sensitive tests. Factors which are included primarily to enhance the generality of results may be relegated to whole plot comparisons. For example, consider again the three-factor experiment of Section 5.2.7, in which permeabilities were compared at $g=4$ levels of gradient (G), $p=2$ levels of permeant (P) and $t=2$ levels of permeameter type (T), with $n=3$ replications of each of the sixteen treatment combinations. If the purpose of the experiment was to test for differences in measured permeability accounted for by permeameter types, and the other two factors were included only in order to increase the scope of the experiment, then a good design would result by dividing the available clay soil into twenty-four batches of size sufficient to accommodate two tests, and assigning each of the eight gradient-permeant combinations at random to three of the batches. Then each batch would be moisterized and divided in two, and the resulting sub-batches assigned at random to the two permeameter types. The order in which the twenty-four batches are processed should be randomized, and within each batch the order in which the tests are set up should also be randomized.

Construction of an analysis of variance table and hypothesis testing procedures in this example are outlined in Appendix B.9. Tests on permeameter type and its interactions should be more sensitive than those obtained from the incomplete block design discussed in Section 5.2.7, since batch size has been reduced from eight to two. The cost of this increased sensitivity is that tests on effects not involving permeameter type will be less sensitive, because they are now based on between-batch comparisons.

Further applied discussion of split-plot designs and their extensions may be found in (34) and (48).

REFERENCES

1. Anderson, T.W. An Introduction to Multivariate Statistical Analysis. John Wiley & Sons, New York, 1958.

2. Bhattacharyya, G.K., and Klotz, J.H. The Bivariate Trend of Lake Mendota. Technical Report No. 98, University of Wisconsin, Madison, Dept. of Statistics, 1966.

3. Bose, R.C., et al., eds. Essays in Probability and Statistics. University of N. Carolina Press, Chapel Hill, NC, 1970.

4. Box, G.E.P. Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems. In: The Annals of Mathematical Statistics, Vol. 25, 1954, pp. 290-302, 484-498.

5. Box, G.E.P, Hunter, W.G., and Hunter, J. S. Statistics for Experimenters. John Wiley & Sons, New York, 1978.

6. Brown, K.W., and Anderson, D.C. Effects of Organic Solvents on the Permeability of Clay Soils. Draft report, U.S. EPA Contract No. R806825010, 1981.

7. Brown, K.W. Data from: A Review and Evaluation of the Influence of Chemicals on the Permeability of Clay Soils. U.S. EPA Contract No. CR-808824, 1984.

8. Bryant, John L. Box-Type Degree of Freedom Adjustments in Crossed Classifications. Dept. of Quantitative Analysis and Information Systems Working Paper, University of Cincinnati, 1985.

9. Bryant, John L. Bayesian Approaches to the Detection of Equilibrium. Dept. of Quantitative Analysis and Information Systems Working Paper, University of Cincinnati, 1985.

10. Cochran, W.G. and Cox, G.M. Experimental Designs, 2nd Ed. John Wiley & Sons, New York, 1957.

11. Company ABC (confidential client identity). Data on laboratory testing on clay soils, 1984.

12. Company XYZ (confidential client identity). Data on laboratory testing on clay soils, 1984.

13. Conover, W.J. Practical Nonparametric Statistics, 2nd Ed. John Wiley and Sons, New York, 1980.

14. Daniel, D.E. Data from: Effects of Hydraulic Gradient and Field Testing on Measurement of Hydraulic Conductivity of Soils. U.S. EPA Contract No. CR-810165, 1984.

15. Das, B.M. Advanced Soil Mechanics. McGraw-Hill, New York, 1983.

16. Degroot, Morris H. Optimal Statistical Decisions. McGraw-Hill, New York, 1970.

17. Dietz, E.J., and Killeen, Timothy J. A Nonparametric Multivariate Test for Monotone Trend with Pharmaceutical Applications. In: Journal of the American Statistical Association, Vol. 76, 1981, pp. 169-174.

18. Dixon, W.J., ed. BMDP Statistical Software. University of California Press, Berkeley, 1983.

19. Dunnett, Charles W. Pairwise Multiple Comparisons in the Unequal Variance Case. In: Journal of the American Statistical Association, Vol. 75, 1980, pp. 796-800.

20. Foreman, David E., and Daniel, David E. Effects of Hydraulic Gradient and Method of Testing of the Hydraulic Conductivity of Compacted Clay to Water, Methanol and Heptane. Draft paper, 1983.

21. Graybill, F.A. Theory and Application of the Linear Model. Duxbury Press, Belmont, CA, 1976.

22. Greenhouse, S.W., and Geisser, S. On Methods in the Analysis of Profile Data. In: Psychometrika, Vol. 32, 1959, pp. 95-112.

23. Hicks, C.R. Fundamental Concepts in the Design of Experiments. Holt, Rinehart and Winston, New York, 1973.

24. Hora, Stephen C., and Conover, W.J. The F-Statistic in the Two-Way Layout with Rank-Score Transformed Data. In: Journal of the American Statistical Association, Vol. 79, 1984, pp. 668-673.

25. Huynh, H., and Feldt, L.S. Estimation of the Box Correction for Degrees of Freedom from Sample Data in the Randomized Block and Split-Plot Designs. In: Journal of Educational Statistics, Vol. 1, 1976, pp. 69-82.

26. Iman, Ronald L., Hora, Stephen C., and Conover, W.J. Comparison of Asymptotically Distribution-Free Procedures for the Analysis of Complete Blocks. In: Journal of the American Statistical Association, Vol. 79, 1984, pp. 674-685.

27. Joyner, Stephenie P., ed. SAS User's Guide: Statistics, Version 5 Ed. SAS Institute Inc., Cary, NC, 1985.

28. Kendall, Maurice G., and Stuart, Alan. The Advanced Theory of Statistics, Vol. 3, 2nd Ed. Hafner Publishing Co., New York, 1968.

29. Krishnaiah, P.R., and Armitage, J.W. Tables for Multivariate t-Distribution. In: Sankhya B, Vol. 28, 1966, pp. 31-56.

30. Levene, H. Robust Tests for Equality of Variance. In: Contributions to Probability and Statistics, I. Olkin, ed. Stanford University Press, Palo Alto, California, 1960. pp. 278-292.

31. Mack, Gregory A., and Skillings, John H. A Freidman-Type Rank Test for Main Effects in a Two-Factor ANOVA. In: Journal of the American Statistical Association, Vol. 75, 1980, pp. 947-951.

32. Mann, Henry B. Nonparametric Tests Against Trend. In: Econometrica, Vol. 13, 1945, pp. 245-259.

33. Miller, Rupert G., Jr. Simultaneous Statistical Inference, 2nd Ed. Springer-Verlag, New York, 1981.

34. Milliken, G.A. and Johnson, D.E. Analysis of Messy Data, Vol. 1: Designed Experiments. Lifetime Learning Publications, Belmont, CA, 1984.

35. Mitchell, J.K., Hooper, D.R., and Campanella, R.G. Permeability of Compacted Clay. In: Journal of the Soil Mechanics and Foundations Division, Vol. 91, No. SM4, 1965. pp. 41-65.

36. Mitchell, J.K. Fundamentals of Soil Behavior. John Wiley & Sons, New York, 1976.

37. Nelson, Wayne. Applied Life Data Analysis. John Wiley & Sons, New York, 1982.

38. Ohio River Division Laboratories, U.S. Army Corps of Engineers, Cincinnati, Ohio. Data on laboratory testing of clay soils, 1984.

39. Olsen, H.W. Hydraulic Flow Through Saturated Clays. Sc.D. Thesis, Massachusetts Institute of Technology, 1961.

40. Peirce, J.J. Data from: Effects of Inorganic Leachates on Clay Liner Permeability. U.S. EPA Contract No. 68-03-3149, 24-1, 1984.

41. Peirce, J. Jeffrey, and Peel, Thomas A. Effects of Inorganic Leachates on Clay Soil Permeability. In: Land Disposal of Hazardous Waste, Proceedings of the Eleventh Annual Research Symposium, U.S. EPA Hazardous Waste Engineering Research Laboratory, Cincinnati, OH, 1985, pp. 182-189.

42. Public Law #98-616. Hazardous and Solid Waste Amendments of 1984, Section 3004, Paragraph "O": Minimum Technological Requirements, November 9, 1984.

43. Resnikoff, G.J., and Lieberman, G.J. Tables of the Non-Central t-Distribution. Stanford University Press, Stanford, CA, 1957.

44. Roberts, D.W. Soil Properties, Classification, and Hydraulic Conductivity Testing. Draft Technical Resource Document for Public Comment, U.S. EPA Publication No. SW-925, 1984.

45. Sheffe, H. The Analysis of Variance. John Wiley & Sons, New York, 1959.

46. Searle, S.R. Linear Models. John Wiley & Sons, New York, 1971.

47. Stoline, Michael R. The Status of Multiple Comparisons: Simultaneous Estimation of All Pairwise Comparisons in One-Way ANOVA Designs. In: The American Statistician, Vol. 35, 1981, pp. 134-141.

48. Winer, B.J. Statistical Principles in Experimental Design, 2nd Ed. McGraw-Hill, New York, 1971.

APPENDIX A

TABLE A-1. TABULATION OF $LOG_{10}(k)$ VERSUS k

$LOG_{10}(k)$	k(cm/sec)
-4.0	1.00×10^{-4}
-4.1	7.94×10^{-5}
-4.2	6.31×10^{-5}
-4.3	5.01×10^{-5}
-4.4	3.98×10^{-5}
-4.5	3.16×10^{-5}
-4.6	2.51×10^{-5}
-4.7	2.00×10^{-5}
-4.8	1.58×10^{-5}
-4.9	1.26×10^{-5}
-5.0	1.00×10^{-5}
-5.1	7.94×10^{-6}
-5.2	6.31×10^{-6}
-5.3	5.01×10^{-6}
-5.4	3.98×10^{-6}
-5.5	3.16×10^{-6}
-5.6	2.51×10^{-6}
-5.7	2.00×10^{-6}
-5.8	1.58×10^{-6}
-5.9	1.26×10^{-6}
-6.0	1.00×10^{-6}
-6.1	7.94×10^{-7}
-6.2	6.31×10^{-7}
-6.3	5.01×10^{-7}
-6.4	3.98×10^{-7}
-6.5	3.16×10^{-7}
-6.6	2.51×10^{-7}
-6.7	2.00×10^{-7}
-6.8	1.58×10^{-7}
-6.9	1.26×10^{-7}

(continued)

TABLE A-1 (continued)

LOG^{10}(k)	k(cm/sec)
-7.0	1.00×10^{-7}
-7.1	7.94×10^{-8}
-7.2	6.31×10^{-8}
-7.3	5.01×10^{-8}
-7.4	3.98×10^{-8}
-7.5	3.16×10^{-8}
-7.6	2.51×10^{-8}
-7.7	2.00×10^{-8}
-7.8	1.58×10^{-8}
-7.9	1.26×10^{-8}
-8.0	1.00×10^{-8}
-8.1	7.94×10^{-9}
-8.2	6.31×10^{-9}
-8.3	5.01×10^{-9}
-8.4	3.98×10^{-9}
-8.5	3.16×10^{-9}
-8.6	2.51×10^{-9}
-8.7	2.00×10^{-9}
-8.8	1.58×10^{-9}
-8.9	1.26×10^{-9}
-9.0	1.00×10^{-9}
-9.1	7.94×10^{-10}
-9.2	6.31×10^{-10}
-9.3	5.01×10^{-10}
-9.4	3.98×10^{-10}
-9.5	3.16×10^{-10}
-9.6	2.51×10^{-10}
-9.7	2.00×10^{-10}
-9.8	1.58×10^{-10}
-9.9	1.26×10^{-10}
-10.0	1.00×10^{-10}

APPENDIX B

STATISTICAL METHODS FOR THE ANALYSIS OF EXPERIMENTAL DATA

B.1 INTRODUCTION

In this Appendix, the construction of analysis of variance tables for the various experimental plans described in Section 5.2 is discussed. In addition, methods for the determination of adequate replication of experiments are given, and diagnostic checks of analysis assumptions are outlined. The final Section of this Appendix describes general considerations concerning the design of a testing program for the evaluation of proposed borrow material to be used in liner construction.

B.2 ANALYSIS OF VARIANCE - COMPLETELY RANDOMIZED DESIGN

Consider the example completely randomized design described in Section 5.2.3, in which k = 4 different chemical permeants are compared in an experiment involving n = 6 replications of each treatment.

Analysis of experimental results is based on the model

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \qquad i=1,2,3,4;\ j=1,2...,6 \tag{1}$$

where y_{ij} is the log-permeability[1] of the j^{th} test of chemical i, μ is a grand mean, τ_i is the effect of the i^{th} chemical ($\Sigma\tau_i=0$) and ε_{ij} is a random error term representing non-systematic fluctuations caused by differences in experimental units and other non-controllable effects. A test of the hypothesis of equal treatment effects (that is, a test of the hypothesis that $\tau_1 = \tau_2 = ... = \tau_k = 0$) is usually based on the analysis of variance table shown in Table B.1. The test uses as a statistic the ratio of the treatment mean square and the error mean square, which is referred to an F-distribution having k-1 = 3 and k(n-1) = 20 degrees of freedom. This test formally requires the assumption that the errors ε_{ij} in equation (1) are normally and independently distributed with common variance σ_ε^2. In Section B.7, tests of these assumptions and remedial measures to be taken if they are violated will be discussed.

---

[1] For purposes of clarity, the discussion in this Section assumes that tests have not been presaturated with standard aquaeous leachate prior to the introduction of chemicals. Whether this should be done is of course a matter of the experimenter's judgement of which method best models the phenomena of interest. If concurrently run tests are all presaturated until water equilibration is achieved, and chemicals are then introduced simultaneously, for some analyses the experimenter may determine that it is more relevant to measure y_{ij} in model (1) relative to its baseline level (see Figure 23). For reasons discussed in Section 4.4 an experiment designed to compare the effects of chemical permeants should generally include water control tests as one of the treatments even if aquaeous presaturation is employed.

TABLE B.1

ANALYSIS OF VARIANCE-COMPLETELY RANDOMIZED DESIGN

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	EXPECTED MEAN SQUARE
Treatments (Chemicals)	SSTR	k-1=3	SSTR/(k-1)	$\sigma_\varepsilon^2+n\sum\tau_i^2/(k-1)$
Error (within Treatments)	SSE	k(n-1)=20	SSE/k(n-1)	σ_ε^2
Total (Corrected)	SSTot	nk-1=23		

Computational formulas for the sums of squares in Table B.1 (and subsequent analysis of variance tables) will not be given here, since in practice these computations would generally be performed by use of appropriate statistical software. For access to these formulas, the reader is referred to any of the texts listed in Section 5.2.1.

If the analysis of variance table B.1 indicates significant treatment differences, then it will usually be of interest to further analyze these differences by means of multiple comparison procedures. Multiple comparison procedures applicable to the designs of Section 5.2 are discussed in (33).

B.3 ANALYSIS OF VARIANCE - RANDOMIZED COMPLETE BLOCK DESIGN

Analysis of the randomized complete block design described in Section 5.2.4 is usually based on the model:

$$y_{ij} = \mu + \tau_i + \beta_j + u_{ij} \quad , \quad i = 1,2,3,4;\ j = 1,2,\ldots,6 \tag{2}$$

where y_{ij} is the log-permeability of the test of chemical i within batch (block) j, μ and τ_i are as defined in equation (1), β_j is the effect of the jth batch, and u_{ij} is a random error term representing non-controllable fluctuations not accounted for by batch differences. A test of the hypothesis of equal treatment effects is based on the analysis of variance table shown in Table B.2 using the ratio of the treatment mean square and the error (treatment x batch interaction) mean square as a test statistic. This ratio is referred to an F-distribution having k-1 = 3 and (n-1)(k-1) = 15 degrees of freedom. The mathematical justification of this procedure requires that the within block errors u_{ij} in model (2) be normally and independently distributed with common variance σ_u^2 (or more generally, that all differences of the form $u_{ij}-u_{\ell j}$ have equal variances).

TABLE B.2

ANALYSIS OF VARIANCE-RANDOMIZED COMPLETE BLOCK DESIGN

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	EXPECTED MEAN SQUARE
Treatments (Chemicals)	SSTR	k-1=3	SSTR/(k-1)	$\sigma_u^2+n\sum\tau_i^2/(k-1)$
Blocks (Batches)	SSB	n-1=5	SSB/(n-1)	$\sigma_u^2+k\sigma_\beta^2$
Error (Treatments x Blocks)	SSE	(k-1)(n-1) = 15	SSE/(k-1)(n-1)	σ_u^2
Total (Corrected)	SSTot	nk-1=23		

B.4 ANALYSIS OF VARIANCE - TWO FACTOR TREATMENT STRUCTURE, COMPLETELY RANDOMIZED DESIGN

The analysis of variance table for the completely randomized factorial experiment described in Section 5.2.5 is shown in Table B.3. A model for the experiment is similar to equation (1):

$$y_{ijk} = \mu + P_i + C_j + PC_{ij} + \varepsilon_{ijk}, \ i = 1,2;\ j = 1,2,3;\ k = 1,2,3,4,5 \qquad (3)$$

in which P_i represents the main effect of permeameter type i (averaged over the four chemicals, $\Sigma P_i = 0$), C_j is the main effect of chemical j ($\Sigma C_j = 0$) and PC_{ij} is the interaction between permeameter i and chemical j ($\sum_i PC_{ij} = \sum_j PC_{ij} = 0$). y_{ijk} represents the log-permeability of the k^th^ replication of permeameter i and chemical j and ε_{ijk} is an unobservable error term accounting for batch differences and other uncontrollable effects. Tests of hypotheses based on Table B.3 formally assume these errors to be independent and normally distributed with common variance σ_ε^2. Tests of any effect (permeameters, chemicals or their interactions) are obtained by computing the ratio of the corresponding mean square to the error mean square and referring this ratio of mean squares to an F-distribution having degrees of freedom as indicated by Table B.3.

In this example, each treatment combination was replicated the same number of times. It is not necessary that this be the case, although in Section B.7 some reasons are given which indicate that balanced replication is to be recommended. However, if treatment combinations are not equally replicated, the sums of squares in Table B.3 are not uniquely defined (46),(34), (27). For most comparative purposes, the most relevant sums of squares in experimental situations such as those discussed here are Type III sums of squares (34), (27), as long as no treatment combination is entirely missing.

TABLE B.3

ANALYSIS OF VARIANCE - TWO WAY TREATMENT STRUCTURE - COMPLETELY RANDOMIZED DESIGN

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	EXPECTED MEAN SQUARE
Treatments				
Permeameters	SSP	p-1=1	SSP/(p-1)	$\sigma_{\varepsilon}^{2}+nc\sum P_{i}^{2}/(p-1)$
Chemicals	SSC	c-1=2	SSC/(c-1)	$\sigma_{\varepsilon}^{2}+np\sum C_{j}^{2}/(c-1)$
Perm x Chem	SSPC	(p-1)(c-1)=2	SSPC/(p-1)(c-1)	$\sigma_{\varepsilon}^{2}+n\sum\sum PC_{ij}^{2}/(p-1)(c-1)$
Error (Within Treatments)	SSE	pc(n-1)=24	SSE/pc(n-1)	σ_{ε}^{2}
Total (Corrected)	SSTot	pcn-1=29		

TABLE B.4

ANALYSIS OF VARIANCE - TWO WAY TREATMENT STRUCTURE REPLICATED IN RANDOMIZED COMPLETE BLOCKS

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	EXPECTED MEAN SQUARE
Treatments				
Permeameter (P)	SSP	p-1=1	SSP/(p-1)	$\sigma_u^2+nc\Sigma P_i^2/(p-1)$
Chemical (C)	SSC	c-1=2	SSC/(c-1)	$\sigma_u^2+np\Sigma C_j^2/(c-1)$
Perm x Chem (PC)	SSPC	(p-1)(c-1)=2	SSPC/(p-1)(c-1)	$\sigma_u^2+n\Sigma\Sigma PC_{ij}^2/(p-1)(c-1)$
Blocks (Batches) (B)	SSB	n-1=4	SSB/(n-1)	$\sigma_u^2+pc\sigma_\beta^2$
Error (Treatments x Blocks)	SSE	(pc-1)(n-1)=20	SSE/(pc-1)(n-1)	σ_u^2
Total (Corrected)	SSTot	pcn-1=29		

B.5 ANALYSIS OF VARIANCE -TWO FACTOR TREATMENT STRUCTURE, RANDOMIZED COMPLETE BLOCK DESIGN

The model usually employed to analyze the two-factor randomized complete block design described in Section 5.2.5 is:

$$y_{ijk} = \mu + P_i + C_j + PC_{ij} + \beta_k + u_{ijk} \qquad i=1,2;\ j=1,2,3;\ k=1,2,3,4,5 \qquad (4)$$

where y_{ijk} is the log-permeability of the sub-batch assigned to permeameter i and chemical j within batch k, and where β_k is the effect of the k^th batch while u_{ijk} models within-batch error. The corresponding analysis of variance table is shown in Table B.4. Tests of the main effect of permeameters, the main effect of chemicals or their interaction are based on F-ratios of the corresponding mean squares to the mean squared error.

B.6 DETERMINING THE REQUIRED NUMBER OF REPLICATES - SPECIFIC METHODS AND EXAMPLES

In Section 5.2.6, two general approaches were discussed for the problem of determining the number of replicates required in an experiment. In this Section, specific details of the methodologies will be outlined, and examples given.

Hypothesis tests of the form discussed in Sections B.2, B.3, B.4 and B.5 employ test statistics which are ratios of mean squares, in which the denominator mean square has expected value σ^2 and the numerator mean square has expected value $\sigma^2 + Q$, where σ^2 is the variance of an appropriate error term, and where Q is a quadratic function of model parameters which is equal to zero if the null hypothesis being tested is true. For example, in the analysis of the completely randomized design of Section 5.2.3 and Appendix B.2, the hypothesis of equal treatment effects is

$$H_0: \tau_1 = \tau_2 = \ldots = \tau_k = 0$$

and is tested using the F-ratio

$$F = MSTR/MSE$$

where the treatment mean square MSTR has, from Table B.1, the expected value $\sigma_\varepsilon^2 + n\sum\tau_i^2/(k-1)$, and the error mean square MSE has expected value σ_ε^2. Thus, in this case, $\sigma^2 = \sigma_\varepsilon^2$ and $Q = n\sum\tau_i^2/(k-1)$. Note that Q = 0 if and only if the null hypothesis H_0 is true.

In general, denote the degrees of freedom of the numerator mean square by f_1, and the degrees of freedom of the denominator mean square by f_2, and suppose that the null hypothesis is to be tested at a level of significance α. Then it can be shown that the probability that the null hypothesis will be rejected (that is, the power of the test) is a function of f_1, f_2, α and the quantity

$$\phi = \sqrt{f_1 Q/(f_1+1)\sigma^2} .$$

Tables of the power as a function of f_1, f_2, α and ϕ are provided in (21), and may be used to compute the power of the hypothesis test against any specified alternative hypothesis.

In designing an experiment, the appropriate number of replicates n may be found by using such tables iteratively. A trial value of n is selected, and the power of the test determined. Then, the value of n is adjusted upward or downward until the required power is obtained. In practice this is usually accomplished within two or three iterations.

To illustrate, consider the planning of an experiment whose purpose is to compare the permeabilities of a clay soil to each of four different dilute chemical permeants. Suppose that it has been decided to use a completely randomized design structure in the experiment, and the required number of replicates n must be determined. The hypothesis of equal permeant effects is

$$H_0: \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$$

and is to be tested at the $\alpha = .01$ level of significance. Suppose it is desired to determine the number of replicates n which is sufficient to yield a .9 probability that H_0 will be rejected if any of the true permeabilities differ by as much as .3 orders of magnitude (or, since $10^{.3} = 2$, by a two to one ratio).

From Table B.1, in this application $Q = n\sum\tau_i^2/(k-1)$, $f_1 = k-1 = 3$, $f_2 = k(n-1) = 4(n-1)$ and $\sigma^2 = \sigma_\varepsilon^2$. Thus

$$\phi = \sqrt{\frac{f_1 Q}{(f_1+1)\sigma_\varepsilon^2}} = \sqrt{\frac{n\sum\tau_i^2}{4\sigma_\varepsilon^2}} .$$

If the range of τ_1, τ_2, τ_3 and τ_4 is .3, then it can be shown that

$$\sum\tau_i^2 \geq (-.15)^2 + 0^2 + 0^2 + (.15)^2 = .045$$

so that ϕ is at least

$$\phi = .106 \sqrt{n} / \sigma_\varepsilon. \qquad (5)$$

At this point, some estimate of the between-test error standard deviation σ_ε is required. In this example the estimate $\sigma_\varepsilon \approx .15$ will be used (as discussed in Section 4.1, the standard deviation of replicated research tests using standard aquaeous leachate or dilute chemicals usually was no greater than this amount in tests included in the data set). This leads to

$$\phi = .71\sqrt{n} .$$

Choosing the trial value $n = 10$ gives $\phi = 2.25$, and $f_2 = 4(10-1) = 36$. Interpolating in Table T.11 of (21) results in a power of about .85. Thus n must be slightly increased. Trying $n = 12$ results in a power of about .92, while $n = 11$ gives a power of about .88. Therefore, in order to obtain a power of at least .9, $n = 12$ replicates of each of the four treatments is required.

To illustrate the comparison of the completely randomized and randomized complete block designs, consider the problem of determining the number of replications required if the experiment were to be replicated in commonly moisturized batches, as in Section 5.2.4. Suppose that it is believed that by comparing treatments within batches, the error standard deviation could be reduced by 33%, so that $\sigma_u \approx .10$. Then from Table B.2, $Q = n\sum\tau_i^2/(k-1)$, $f_1 = k-1 = 3$ and $f_2 = (n-1)(k-1) = 3(n-1)$. From equation (5)

$$\phi = .106\sqrt{n}/.10 = 1.06\sqrt{n} .$$

Proceeding as above, Table T.11 of (21) indicates that $n = 7$ batches are required to obtain a power of at least .9. This example illustrates that if blocking of experimental units is even moderately effective in reducing between-test variability, the economy that can be obtained by use of replication in complete blocks may be substantial.

An alternative method of sample size determination is available if the experimenter has apriori a particular interest in certain specific treatment comparisons. If this is the case, the number of replications may be determined such that the widths of confidence intervals for these comparisons do not exceed certain specified bounds. For example, in the experiment considered in this Section, one of the four permeants to be tested would generally be a control such as standard aquaeous leachate or distilled water. In this situation, there are three treatment comparisons that are probably of primary interest to the experimenter. These are the comparisons of the three "active" chemicals with the control. A set of joint 99% confidence intervals for these comparisons is given by Dunnett's method as

$$(\bar{y}_i - \bar{y}_1) \pm d_{.005,f_2}\sqrt{\frac{2\ \mathrm{MSE}}{n}} \ , \ i = 2,3,4$$

where $\bar{y}_1$ is the mean log permeability under the control, and $\bar{y}_2$, $\bar{y}_3$ and $\bar{y}_4$ are the mean log permeabilities for the remaining three chemicals. $d_{.005,f2}$ is Dunnett's two sided critical value based on f_2 degrees of freedom (which is also dependent on the number of treatments in the experiment, $k = 4$ in this example). Such critical values are tabulated in (48).

Suppose that it is desired to replicate the experiment a sufficient number of times to estimate the effects of the active chemicals relative to the control to within plus or minus .2 orders of magnitude, with 99% confidence. This would require

$$d_{.005,f_2}\sqrt{\frac{2MSE}{n}} = .2 \ .$$

Since MSE is not observable until the experiment is complete, a prior estimate of the error variance is needed. If the experiment is to be replicated in complete blocks, and if the same estimate $\sigma_u^2 \approx (.10)^2$ is used that was employed in the earlier power calculations, then the required number of replications may be estimated by solving

$$d_{.005,f_2}\sqrt{\frac{2(.10)^2}{n}} = .2 \qquad (6)$$

for n, where $f_2 = (k-1)(n-1) = 3(n-1)$ is the degrees of freedom for error, as given in Table B.2. Equation (6) may be solved iteratively until the least value of n is found such that the left-hand side of the expression is no more than .2. Using Table C.6 in (48), it may be seen that the required number of replications is n = 6.

Either of the methods of sample size determination which have been reviewed in this Section require some estimate of the error standard deviation. An unfortunate practical difficulty is that the amount of replication called for by either method is quite sensitive to the estimate which is used. Whenever possible, such estimates should be based on the results of previous investigations which have made use of soils, permeants and permeameters similar to those which will be employed in the experiment to be designed. Lacking other sources of information, the following rule of thumb may be useful in obtaining a rough benchmark: If the experimenter believes that a very large number of replications of any of the treatments (under conditions similar to those to be found in the experiment) would result in a range of R orders of magnitude in resulting permeabilities, then estimate the error standard deviation by $\sigma = R/4$.

B.7 DIAGNOSTIC CHECKS AND REMEDIAL MEASURES

The methods of analysis described in Sections B.2, B.3, B.4 and B.5 are based on certain assumptions which concern the distribution of error terms in model equations (1), (2), (3), and (4). In this Section a brief discussion of these assumptions is given, with consideration of the sensitivity of the methods to violations of the assumptions, procedures to identify such violations if they occur, and remedial measures which may be taken if warranted.

Normality: All methods of hypothesis testing or confidence interval construction discussed in Sections B.2, B.3, B.4 and B.5 are based on an assumption of normally distributed errors. Fortunately, moderate non-normality of errors does not appear to have a serious impact on the nominal levels of significance (or confidence) of these procedures (5).

It is highly recommended that after any model is fit to the data, plots of residuals be constructed and inspected. Useful plots include stem-and-leaf plots of residuals, plots of residuals against treatments and blocks, against predicted values, against the order in which tests are set up, and more generally against any variable which theory or practice might suggest to the experimenter. A straightforward description of the interpretation of such plots is found in (5). Plotting of residuals is easily accomplished when using any of the standard computer data analysis packages such as SAS, BMDP or SPSSX, either by invoking options within the procedures used in the analysis of variance of the data, or by outputting the residuals to a new data set which then may be used as input to specialized plotting or data summarization procedures. For example, if SAS is used, residuals from either PROC ANOVA or PROC GLM may be output to PROC UNIVARIATE for construction of a stem-and-leaf plot and goodness-of-fit tests for normality (either the Shapiro-Wilk or Kolmogorov-Smirnov test), and to PROC PLOT for the remainder of the plots described above.

Visual inspection of plots will indicate serious deviations from normality, and are supported by formal goodness-of-fit tests as described above. Such tests are in general only approximate, since the residuals are subject to certain linear restrictions. The effect of these restrictions should be relatively minor if the degrees of freedom for error are considerably larger than those for treatments and blocks, but if this is not the case, some spurious appearance of non-normality may result. Residual plots are also useful in identifying non-constant error variance and the occurrence of very deviant responses, which should be investigated for circumstances which may have caused an unusual result.

Non-normality of errors may sometimes be corrected by an appropriate transformation of the data (5). For example, one reason for the recommendation made in Sections B.2, B.3, B.4, and B.5 that log-permeabilities be analyzed rather than raw permeabilities is that empirical evidence indicates that the data will appear more normal when expressed on a logarithmic scale (see also Section 4.1).

If residual analysis indicates marked non-normality of residuals, use of nonparametric procedures, which do not require a normality assumption, are recommended. At the very least, comparison of nonparametric results with those of the usual parametric analysis of variance is a good way to judge the sensitivity of conclusions to the normality assumption.

The usual nonparametric alternative in the one-way completely randomized case discussed in Section 5.2.3 and Appendix B.2 is the Kruskal-Wallis test, which essentially amounts to an analysis of variance on ranks. This procedure is discussed in (13).

In the randomized complete block situation of Section 5.2.4 and Appendix B.3, there are several nonparametric alternatives in current use. These include the Freidman procedure, Quade's test, and the analysis of variance of total-sample ranks. A comparison of these three methods is made in (26), and a more general discussion of the procedures is found in (13).

References (24), (31) and (13) discuss nonparametric and rank-transform methods for the analysis of factorial experiments, and list additional references.

Independence: Analyses of experimental data from experiments with completely randomized design structures, which were outlined in Appendices B.2 and B.4, assume that responses may be treated as independent random variables. In practice, these responses may be correlated for a number of reasons. For example, gradual improvement in compaction or other preparation techniques during the course of setting up tests may result in a positive correlation of results of tests which were set up at nearly the same time.

The effects of such correlation may seriously invalidate the results of formal statistical tests of hypotheses or methods of confidence interval construction. Fortunately, the experimenter has control of this situation not only through the direct control of extraneous factors which might influence experimental results, but also through the randomization of the assignment of treatments to experimental units and the order in which tests are set up. If proper randomization of the experiment is included in its design, the effects of extraneous factors which might otherwise bias treatment comparisons will be mitigated.

Analogous statements may be made concerning the analysis of experiments which are replicated in randomized complete blocks. Here, the responses to treatment combinations compared within the same block may be correlated (by virtue of the fact that they have been applied to experimental units from within the same block), but it is assumed that the correlation between any pair of responses is the same (or slightly more generally, that the variances and covariances of these responses obey a set of restrictions known as the Huyhn-Feldt conditions). Such an assumption would generally be completely unwarranted unless care is taken to ensure proper within-block randomization as discussed in Sections 5.2.4 and 5.2.5.

Nonconstant Variance: The usual analysis of variance of data from experiments having completely randomized design structures assumes that the variances of responses to each treatment combination is the same. Thus, for example, the analysis of the 2 x 3 factorial experiment based on Table B.3 in Appendix B.4 assumes that variation of log-permeabilities is the same for each of the six permeameter-chemical treatment combinations included in the treatment structure of that experiment.

In permeability experimentation, it is reasonable to expect that the assumption of constant variance may often be violated. This problem may sometimes be corrected by data transformation. For example, the discussion in Section 4.1 indicated that use of the logarithmic transformation tends to promote the homoscedasticity (constant variance) of permeability measurements over different clay soils. However, even on the logarithmic scale heteroscedasticity may occur. As was noted in Section 4.1, the variability of log-permeabilities for different clay-permeant combinations seemed to be rather dependent on the particular combination investigated, and on the basis of very limited data, the use of rigid-wall permeameters in some cases seemed to yield more variable results than did flexible-wall permeameters, as discussed in Section 4.6.

A large number of diagnostic procedures have been developed to test for constant variance, including Bartlett's test, Cochran's test, the F-max procedure, Box's test and Levene's test (34), (48), (30). In general, our preference is to use Levene's test, because it is highly robust to non-normality. Levene's test is available in the BMDP (18) data analysis package (program BMDP7D).

Fortunately, there is both theoretical and empirical support for the fact that the usual hypothesis tests provided by analysis of variance are quite robust to violations of the constant variance assumption if each treatment combination is replicated the same number of times. On the other hand, if cell sample sizes are noticably uneven, the usual F-tests can be quite sensitive to this assumption. These facts provide considerable motivation to plan permeability experiments to include equal replication of all treatment combinations of interest.

In the case of a completely randomized design having a one-way treatment structure with k treatments and n replications, Box (4) has shown that the F-test for the comparison of treatments has a distribution which is well-approximated by that of an F-distribution with $\varepsilon(k-1)$ numerator and $\varepsilon'k(n-1)$ denominator degrees of freedom, where

$$\varepsilon = \frac{1}{1 + \frac{k-2}{k-1} c^2} , \qquad (7)$$

$$\varepsilon' = \frac{1}{1 + c^2}$$

and where c is the coefficient of variation of the k treatment variances (in the case where these variances are equal, $c^2=0$ so that $\varepsilon=\varepsilon'=1$, and the results reduce to those summarized in Table B.1). This result has led Milliken and Johnson (34) to suggest the following procedure: Prior to performing the analysis of variance, test for homoscedasticity using Levene's test. If the hypothesis of equal variances cannot be rejected at the 1% level of significance, perform the usual analysis of variance as summarized in Table B.1. If the hypothesis of homoscedasticity is rejected, then estimate the correction factors ε and ε' by inserting the coefficient of variation of the k sample variances into equations (7), and adjust the degrees of freedom used in the F-test accordingly.

The procedure described above may be extended to two-way or higher-way factorial experiments by making use of correction factors similar to those in equation (7) which are derived and given in (8).

Use of the above results requires an equal number of replications for each treatment combination. In situations where unequal replication is unavoidable, Milliken and Johnson (34) recommend the use of Welsh's test, which is essentially a weighted least squares procedure. Welsh's test is available (for the one-way classification) in BMDP program BMDP7D.

In the presence of heteroscedasticity, the use of the overall mean squared error in construction of confidence intervals for specified individual treatment comparisons, or in tests of hypotheses concerning such comparisons, may lead to erroneous results. Appropriate techniques in such cases are based on the Satterthwaite approximation, and are discussed in (34). Methods of multiple comparison in the unequal variance case are described in (19).

In experiments which are replicated in commonly moisturized batches, in accordance with a randomized complete block design structure, the homogeneity of response variances is also an issue. As mentioned above, the usual analyses require a set of conditions, known as the Huynh-Feldt conditions. These conditions are equivalent to supposing that the differences of responses to every pair of treatments within a batch have the same variance. While this is mathematically possible in cases where treatment response variances are not all equal, in practice heteroscedasticity would almost certainly imply a violation of the Huynh-Feldt conditions. An exception to this statement occurs in randomized complete block experiments in which only two treatments are compared, since then there is only one treatment difference to consider.

Results due to Box (4) indicate that the use of the usual F-test is rather insensitive to nonconstant variance in the randomized complete block design, but is somewhat more sensitive to departures from the Huynh-Feldt conditions which may be caused by non-symmetric covariation between treatment responses. These latter effects may be minimized by proper within-block randomization, as discussed above.

A test for departure from the Huynh-Feldt conditions (applicable in cases where the number of batches is large) has been developed by adapting Anderson's sphericity test (1) to orthogonal contrasts of treatment responses, and has been incorporated into BMPD program BMDP2V and SAS procedures ANOVA and GLM (version 5). Degree of freedom adjustments due to Greenhouse and Geisser (22) and to Huynh and Feldt (25) have also been incorporated into these programs, and provide approximately valid tests in cases where the Huynh-Feldt conditions do not hold.

Single degree of freedom treatment comparisons may be analyzed in randomized complete block designs without any theoretical requirement of homoscedasticity. Consider, as an example, an experiment in which the effects of four permeants - a control such as standard aquaeous leachate, and three "active" chemicals -are to be compared by preparing n moisterized batches of a clay soil, randomly quartering each batch, and applying each permeant to one sub-batch within each batch. Then, each of the three active chemical versus control contrasts is a single degree of freedom comparison, which may be tested by ignoring the remaining two permeants and treating the resulting data as though it had come from a randomized complete block design comparing just one of the active chemicals with the control. The resulting analysis is equivalent to a paired t-test which requires no equal variance assumption. An overall level of significance for the three comparisons may be maintained, if desired, by using Bonferroni's inequality (33). For example, if a joint significance of .05 were desired for tests of the three treatment versus control comparisons, each one could be individually tested at the .05/3 = .017 level.

B.8 ANALYSIS OF VARIANCE - INCOMPLETE BLOCK DESIGN WITH PARTIAL CONFOUNDING OF FACTORIAL EFFECTS

In Section 5.2.7, a 2x2x4 factorial experiment investigating the factors P (permeant, at p = 2 levels), T (permeameter type, at t = 2 levels) and G (gradient, at g = 4 levels) was considered. Each of the 16 treatment combinations of the experiment was replicated three times, in six incomplete blocks (commonly moisturized batches) of eight tests, as shown in Table 4. This was done in such a way that all main effects and two factor interactions are free of batch differences. However, the three degrees of freedom for the PTG interaction are partially confounded, with one degree of freedom confounded with batch differences in each of the three replications of the experiment. Specifically, if the interaction of P and T at level j of factor G is denoted by η_j, $j = 0,1,2,3$, then it may be shown that

$\eta_0 - \eta_1 - \eta_2 + \eta_3$ is confounded with batch differences in replication #1.

$\eta_0 - \eta_1 + \eta_2 - \eta_3$ is confounded with batch differences in replication #2.

$\eta_0 + \eta_1 - \eta_2 - \eta_3$ is confounded with batch differences in replication #3.

The general form of the analysis of variance table for this experiment is shown in Table B.5. Mean squares are obtained by dividing sums of squares by the degrees of freedom indicated in the table. Tests of treatment effects are based on F-ratios obtained by dividing the corresponding mean squares by the error mean square. Generally the sums of squares for the three single degree of freedom components of the three-factor interaction are pooled as shown in Table B.5 prior to testing.

TABLE B.5

ANALYSIS OF VARIANCE TABLE
INCOMPLETE BLOCK DESIGN

SOURCE	DEGREES OF FREEDOM	
Between Batches		
Between Replications	2	
Batches Within Replicates	3	
Within Batches		
P	1	
T	1	
PT	1	
G	3	
PG	3	
TG	3	
PTG		
Component PTG_1	1	(from Replicates 2 and 3)
Component PTG_2	1	(from Replicates 1 and 3)
Component PTG_3	1	(from Replicates 1 and 2)
Error	27	(by subtraction)
Total (Corrected)	47	

For computational purposes, the easiest way to compute the required sum of squares and mean squares is to fit a model of the form

$$y_{ijk\ell} = \mu + P_i + T_j + PT_{ij} + G_k + PG_{ik} + TG_{jk} + PTG_{ijk} + \beta_\ell + u_{ijk\ell} \quad (8)$$

to the data by means of an appropriate statistical software package. In equation (8) $y_{ijk\ell}$ is the log-permeability obtained in the test within batch ℓ using permeant i, permeameter type j and gradient k (ℓ = 1,2,3,4,5,6; note that only 48 of the 96 possible combinations of i,j,k and ℓ appear in the data). P_i, T_j and G_k represent main effects, PT_{ij}, PG_{ik} and TG_{jk} are two-factor interactions. PTG_{ijk} is the three factor interaction, β_ℓ is the effect of the ℓth batch, and $u_{ijk\ell}$ is a random error term. The sum of squares obtained for the main effects and two-factor interactions when model (8) is fit are not dependent on the presence or absence of the batch effects β_ℓ, but this is not the case for the three-factor interaction PTG. Accordingly, the Type III mean square for PTG should be used for purposes of testing for the presence of the three-factor interaction.

B.9 ANALYSIS OF VARIANCE - SPLIT PLOT DESIGNS

Section 5.2.8 contains two examples of split-plot permeability experiments. An analysis of variance table for the first example is shown in Table B.6. The general notation used in that table is as follows: w represents the number of levels of the whole plot factor molding water content (w=2 in this example), g is the number of levels of the sub-plot factor gradient (g=3) and n stands for the number of times each treatment combination is replicated (n = 5). Expected mean squares are derived under the model

$$y_{ijk} = \mu + W_i + \varepsilon_{ij} + G_k + WG_{ik} + u_{ijk}$$

$$i=1,2,\ldots,w; \quad j=1,2,\ldots,n; \quad k=1,2,\ldots,g$$

where y_{ijk} represents the log-permeability of the sub-batch from batch j which was compacted at water content i and tested at gradient k, W_i, G_k and WG_{ik} are the main effects and interactions of molding water content and gradient, ε_{ij} is a random error term representing batch-to-batch variability (whole plot error) and u_{ijk} is a random error term which models sub-batch variability within batches (sub-plot error). These error terms are assumed to be normally and independently distributed with variances σ_ε^2 and σ_u^2 respectively.

Inspection of the expected mean squares in Table B.6 shows that a test of the effect of molding water content should be based on the F-ratio of MSW=SSW/(w-1) to the whole-plot mean squared error MSB(W)=SSB(W)/w(n-1), whereas tests of the gradient effect or the water content-gradient interaction should be based on ratios of their mean squares to the sub-plot mean squared error MSBG(W)=SSBG(W)/w(n-1)(g-1). The expressions for expected mean squares also indicate that tests of the within-plot comparisons should be considerably more sensitive than the test for differences due to molding water content. This should not be surprising, since the experiment was arranged in such a way that gradient effects were compared under more uniform conditions than were molding water content effects. This is not a serious concern, since the primary purpose of the experiment was the comparison of gradient effects, with

TABLE B.6

ANALYSIS OF VARIANCE TABLE: SPLIT-PLOT DESIGN

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	EXPECTED MEAN SQUARE
Between Batches				
Water Content	SSW	w-1=1	SSW/(w-1)	$\sigma_u^2 + g\sigma_\varepsilon^2 + ng\Sigma W_i^2/(w-1)$
Batches Within Water Content (Whole Plot Error)	SSB(W)	w(n-1)=8	SSB(W)/w(n-1)	$\sigma_u^2 + g\sigma_\varepsilon^2$
Within Batches				
Gradient	SSG	g-1=2	SSG/(g-1)	$\sigma_u^2 + nw\Sigma G_k^2/(g-1)$
Water Content x Gradient	SSWG	(w-1)(g-1)=2	SSWG/(w-1)(g-1)	$\sigma_u^2 + n\Sigma\Sigma WG_{ik}^2/(w-1)(g-1)$
Batches x Gradient Within Water Content (Sub-Plot Error)	SSBG(W)	w(n-1)(g-1)=16	SSBG(W)/w(n-1)(g-1)	$\sigma^2{}_u$
Total (Corrected)	SSTot	wng-1=29		

the water content factor considered primarily for purposes of enhancing the generalizability of the experimental results. In addition, the effect of a 5% difference in molding water content is liable to be quite pronounced, so that extreme sensitivity is probably not required.

The methods of Section B.6 may be applied to the analysis of variance Table B.6 to determine the number of replications required for an experiment such as that described above. For example, n could be determined by specifying the desired power of the test of gradient effects under an appropriately specified alternative hypothesis.

Diagnostic checks of data from split-plot designs follow generally along the lines discussed in Section B.7. The analysis of the whole plot factor(s) may be checked by summing or averaging over all levels of the sub-plot factor(s), resulting in one score for each plot. These summarized data may then be checked as though they had resulted from an experiment with a completely randomized design structure, and if required, remedial action may be taken as described in Section B.7.

The within-plot analyses require assumptions analogous to those for experiments which are replicated in randomized complete blocks. Responses from different batches are taken to be independent, and within-batch variances and covariances are assumed to obey the Huynh-Feldt conditions. This may be tested by use of Anderson's sphericity test, available in BMDP program BMDP2V and SAS Procedures ANOVA and GLM. These software packages also provide degree of freedom adjustments appropriate for approximate tests if the conditions are violated, as well as tests on single degree of freedom contrasts of within-plot treatments, which do not require the Huynh-Feldt conditions for their validity.

An assumption of identical within-batch variance-covariance matrices (across all levels of the whole plot factors) is also implied by the usual analysis. This assumption may be checked by use of a test due to Box (48).

An analysis of variance table for the second split-plot experiment of Section 5.2.8 is given in Table B.7. Expected mean squares indicate that the sub-plot error should be used to form tests of permeameter type and all interactions involving permeameter type, whereas the whole plot error is used in tests involving only gradient and/or permeant effects.

Diagnostic checks for the experiment of Table B.7 are straightforward. As mentioned above, the analysis of the whole-plot factors is equivalent to a completely randomized design analysis of variance on scores obtained by summing over the two permeameter types, and may be checked accordingly. In addition, since in this example there are only two permeameter types to be compared, the Huynh-Feldt conditions are automatically satisfied. In this case, the sub-plot analysis is equivalent to a completely randomized design analysis of variance on scores obtained by differencing over the two permeameter types, which leads to straightforward methods of diagnostic checking and (if appropriate) remedial action.

TABLE B.7

SPLIT-PLOT DESIGN WITH TWO WHOLE PLOT FACTORS AND ONE SUB-PLOT FACTOR

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	EXPECTED MEAN SQUARE
Between Batches				
Gradient G	SSG	g-1=3	SSG/(g-1)	$\sigma^2_u + t\sigma^2_\epsilon + ntp\Sigma G^2_i/(g-1)$
Permeant P	SSP	p-1=1	SSP/(p-1)	$\sigma^2_u + t\sigma^2_\epsilon + ntg\Sigma P^2_j/(p-1)$
Gradient x Permeant GP	SSGP	(g-1)(p-1)=3	SSGP/(g-1)(p-1)	$\sigma^2_u + t\sigma^2_\epsilon + nt\Sigma\Sigma GP^2_{ij}/(g-1)(p-1)$
Batches Within GP (Whole Plot Error)	SSB(GP)	pg(n-1)=16	SSB(GP)/pg(n-1)	$\sigma^2_u + t\sigma^2_\epsilon$
Within Batches				
Permeameter Type T	SST	t-1=1	SST/(t-1)	$\sigma^2_u + npg\Sigma T^2_k/(t-1)$
GT	SSGT	(g-1)(t-1)=3	SSGT/(g-1)(t-1)	$\sigma^2_u + np\Sigma\Sigma GT^2_{ik}/(g-1)(t-1)$
PT	SSPT	(p-1)(t-1)=1	SSPT/(p-1)(t-1)	$\sigma^2_u + ng\Sigma\Sigma PT^2_{jk}/(p-1)(t-1)$
GPT	SSGPT	(p-1)(g-1)(t-1)=3	SSGPT/(p-1)(g-1)(t-1)	$\sigma^2_u + n\Sigma\Sigma\Sigma GPT^2_{ijk}/(p-1)(t-1)(g-1)$
Batches x Type Within GP (Sub-Plot Error)	SSBT(GP)	pg(n-1)(t-1)=16	SSBT(GP)/pg(n-1)(t-1)	σ^2_u
Total (Corrected)	SSTot	gptn-1 = 47		

B.10 PERMEABILITY TESTS OF PROPOSED BORROW MATERIAL

In this Section, general considerations concerning the design of a testing program for the evaluation of borrow material, which is under consideration for use in the construction of a liner, are discussed. The situation here differs from that described in Section 5 primarily in that here the major goal of the investigation is the estimation of the permeability of the soil within the proposed site, in order to determine whether it will meet specified performance criteria. In Section 5, on the other hand, the experiments which were described were of a comparative nature, and were intended to investigate possible differences caused by various treatments, such as different permeameter types or different chemical permeants.

One consideration which is important in this discussion is that the soil contained within the proposed site may exhibit a good deal of heterogeneity in its chemical and physical properties. This may lead to considerable variation in permeability (see Section 4.2) so that an estimate of the average permeability throughout the site may not be very relevant in judging its overall acceptability. That is, a source whose average permeability might be considered to be acceptably small for use as liner material might nevertheless contain a significant amount of material having substandard permeability characteristics.

For this reason, it would be prudent to partition the site into smaller sub-sites, which will be referred to as "cells", in such a way that the characteristics of the soil found within any one cell are reasonably uniform. This could be done based on a preliminary determination of characteristics such as Atterberg limits, grain size distribution and water content-density relationships. Such cells should not be made up of noncontiguous parts, should include a single soil layer, and at a minimum should contain soils having the same Unified Soil Classification.

Once the site has been stratified into cells, replicated permeability tests can be made on samples drawn from different locations within each cell. This will allow an estimation of the permeability of the soil within each cell, which may be used to determine which, if any, of the cells are not suitable for use as liner material.

In the following, it will be supposed that the site has been partitioned into c cells as described above, and that samples are available from n_i randomly selected locations within the ith cell, $i=1,2,\ldots,c$. Random selection of locations within a cell is facilitated by laying a grid over a map of the cell, followed by random selection of the required number of squares lying within the boundaries of the cell to determine sampling locations. For each selected location, the depth at which the sample is to be taken should also be randomly selected within the depth limits of the cell at that point. (As a matter of practical economy, any suitable soil samples available from the preliminary site investigation might also be used for permeability testing.)

The $N = \sum_{i=1}^{c} n_i$ permeability tests are set up in random order, following standardized methods of specimen preparation, compaction and testing. Let y_{ij}

denote the log-permeability of the sample from location j within cell i, $i=1,2,\ldots,c$; $j=1,2,\ldots,n_i$. Then $\overline{y}_i = \sum_j y_{ij}/n_i$ gives an estimate of μ_i, the mean log-permeability of the soil contained in cell i.

If the distribution of log-permeabilities within cells is at least approximately normal, and the within-cell variability of log-permeabilities is constant across cells, then a set of confidence intervals for $\mu_1, \mu_2, \ldots, \mu_c$ having joint confidence $(1-\gamma)$x100% is given by

$$\overline{y}_i \pm t_{\gamma/2,\nu,c} \sqrt{\frac{MSE}{n_i}} \text{ , } i=1,2,\ldots,c \tag{9}$$

where $t_{\gamma/2,\nu,c}$ is the $1-\gamma$ fractile of the distribution of the maximum absolute value of c studentized normal variates based on $\nu=N-c$ degrees of freedom, and where MSE is the pooled within-cell variance of the y_{ij}'s. Tables of the critical values $t_{\gamma/2,\nu,c}$ are provided in (21). In equation (9), substituting the $(1-\gamma/2c)$th fractile of Student's t-distribution in place of $t_{\gamma/2,\nu,c}$ gives, by Bonferroni's inequality, an alternate method for computing these intervals. Bonferroni intervals will be somewhat wider than those given in (9), but eliminate the need for specialized tables.

Statistical tests appropriate to form a basis on which to judge the acceptability of the soil depend on how current requirements are interpreted. For example, one interpretation might be that the soil within cell i is to be judged acceptable if and only if the null hypothesis that the median permeability within cell i is at least 10^{-7} cm/sec can be rejected against the alternative hypothesis that it is less than 10^{-7} cm/sec, at some specified level of significance. For symmetrically distributed log-permeabilities, these hypotheses are equivalent to:

$$\begin{aligned} H_i&: \mu_i \geq -7 \\ H_i'&: \mu_i < -7, \ i = 1,2,\ldots,c \ . \end{aligned} \tag{10}$$

The hypotheses in (10) may be tested by rejecting H_i in favor of H_i' if and only if

$$\frac{\overline{y}_i - (-7)}{\sqrt{MSE/n_i}} < -t_{crit} \ . \tag{11}$$

Choice of the critical value t_{crit} in (11) depends on which error rate is to be controlled. Using $t_{\gamma,\nu}$ the $(1-\gamma)$th fractile of Student's t-distribution, will result in c tests, each of which has an individual level of significance of γ. If $t_{\gamma/c,\nu}$ is used, the joint probability that any cells will be erroneously judged acceptable will be controlled to be no more than γ, by Bonferroni's inequality. A somewhat more powerful test which controls the joint level of significance is obtained by basing t_{crit} on the distribution of the maximum of c studentized normal variates. Required tables may be found in (29) and (3).

Up to this point no mention has been given to the selection of the sample sizes $n_1, n_2, \ldots, n_c$. Generally it would be expected that these would not be chosen to be equal, since knowledge of the general characteristics of the soils within the c cells would allow an experienced technician to form reasonable apriori estimates of their permeabilities. Then, it would make sense to sample more frequently from the cells whose soil is thought to have permeability close to the 10^{-7} cm/sec requirement, and less frequently from the cells which are predicted to be quite impermeable. A method for choosing appropriate sample sizes $n_1, n_2, \ldots, n_c$ consists of making prior estimates μ_i^* for the log-permeability of each cell i, then determining n_i so that the hypothesis test (11) would result in the rejection of the null hypothesis H_i in (10) with a specified probability $1-\beta$ if in fact μ_i were equal to μ_i^*. This may be done in a manner similar to Appendix B.6, which is described in (48). This reference also provides specialized tables to facilitate the sample size selection process. More complete tables are found in (43). If desired, the joint power of all c tests can be controlled by Bonferroni's inequality.

As in Appendix B.6, some prior estimate of the within-cell standard deviation of log-permeabilities will be required in the determination of sample sizes. If no other source of obtaining such an estimate is available, the numerical results given in Section 4.2 might be used to provide a rough benchmark, although for reasons discussed in that Section their generalizability is highly suspect.

The hypothesis testing scheme outlined above is reasonable if current EPA permeability requirements are interpreted in terms of median cell permeabilities, but other interpretations could be made as well. For example, one could test to see whether a specified percentage p% of the soil within each cell had a permeability beneath the 10^{-7} cm/sec cut-off. Such tests would be based on the same decision rule shown in (11), except that in this case the critical value is based on the non-central t-distribution. A more complete description of the test procedure, and tables of the critical values required for the test, may be found in (43).

In the testing program which has been outlined in this Section, tests are conducted under specified conditions of molding water content and compactive effort. If the test results are to be representative of the future performance of the proposed material, then presumably these test conditions would be determined to be the same as the conditions obtaining when the actual liner is constructed. If molding water content and compactive effort can be closely controlled during the construction of a liner, then this should be straightforward. On the other hand, if during construction it is to be expected that these parameters may vary significantly about their specified target values, then the results of tests conducted at these target values may misrepresent the future performance of the liner, since the permeability of a soil is strongly related to both compactive effort and molding water content (see (35), (44) and Section 4.3).

If variation in compactive effort and molding water content during construction is significant, then the design of an adequate permeability testing program would require an understanding of the bounds within which these parameters can be maintained, so that tests may be conducted to determine whether the proposed liner material will be acceptably impermeable under all conditions within these bounds. For example, under the assumption that permeability is a decreasing function of compactive effort, tests should be conducted at the minimum compactive effort that could be guaranteed to be maintained during liner application.

With respect to molding water content, the situation is not as clear-cut. The results in Section 4.3 indicate that permeability is a decreasing function of molding water content in the general vicinity of the optimum water content (for density), and that the point at which permeability begins to increase with molding water content occurs several percent above optimum. (This phenomenon has been noted as early as 1965 (35); however, caution should be applied in assuming it will generalize in all circumstances. It is conceivable that for certain permeants other than water the mechanisms which lead to decreasing permeability even after the point at which density begins to decrease may be dominated by other considerations.)

In some circumstances, depending on the target molding water content and the tightness of the bounds which can be maintained during liner application, an investigator may be willing to assume that tests conducted at the lower bound would provide an adequate estimation of the "worst case" performance of the proposed material. However, if this is not the case, then it would be prudent to alter the testing program outlined in this Section to include tests from each cell at both the upper and lower bounds on molding water content. An assessment could then be made, essentially as above, to determine whether the proposed material is acceptable at either extreme condition. If this were the case, then given the general shape of the permeability-molding water content curves shown in Section 4.3, it would be reasonable to assume that the material would be acceptable at any intermediate value of molding water content.

APPENDIX C

METHODS TO CONFIRM EQUILIBRIUM OF RESEARCH PERMEABILITY TESTS

Because the measured permeability of a test often undergoes a prolonged period of transitory behavior before leveling off at an equilibrium value, deciding when the test has exhibited stable permeability long enough to conclude that equilibrium has been achieved is not straightforward. Development of objective criteria, which may be useful in confirming the experimenter's judgment in this decision, is a topic which deserves considerable attention in future permeability research. In this Section, two approaches are described which are currently under investigation. The first is an adaptation of a test for trend originally proposed by Mann (32), which has been heuristically modified for use as a test termination rule in a manner similar to the termination procedure developed by Peirce (41). The second approach is based on a Bayesian analysis (16) in which the posterior probability that equilibrium has been achieved is updated after each new observation, until this posterior probability first falls below a specified bound.

Before describing these methods, three points should be emphasized. The first is that neither of the methods are proposed as the "best" solution to the problem of detecting the point at which equilibration occurs. Rather, in the authors' opinion they represent reasonable approaches to the problem which are deserving of additional consideration and empirical testing. They are described here in the hope that they will stimulate discussion of the problem and may form a framework upon which future research may build.

The second point is that tests of the nature described here will probably be of practical value only in situations where tests approach their equilibrium levels in a relatively continuous manner. They are not intended to be able to predict the behavior of tests which are subject to sudden breakthroughs leading to episodic increases in permeability.

Finally, it should be emphasized that methods such as those described here should not be regarded as replacing the experimenter's informed judgment. Rather, they are tools which are potentially useful in providing objective confirmation of the experimenter's judgment in determining that a test has equilibrated. In no case should an experimenter permit his or her judgment to be overridden in a situation where he or she feels that a test has not equilibrated.

Mann's Test for Monotone Trend:

Mann (32) has described the following test for time trend of a sequence $y_1, y_2, \ldots, y_n$ of continuous observations, ordered over time. Define

$$\text{sign}(x) = \begin{cases} +1, & \text{if } x > 0 \\ 0, & \text{if } x = 0 \\ -1, & \text{if } x < 0 \end{cases}$$

Then the test statistic is

$$K = \sum_{1 \leq i < j \leq n} \text{sign}\,(y_j - y_i).$$

Under a null hypothesis of random ordering of the y_j's, K has an expected value of 0 and a variance of $n(n-1)(2n+5)/18$. If n is 10 or more, the distribution of K is well approximated by a normal distribution. Thus, to test the hypothesis of random ordering against an alternative hypothesis of increasing time trend, the standardized statistic $K/\sqrt{n(n-1)(2n+5)/18}$ is compared to the $(1-\gamma) \times 100\%$ percentile of the standard normal distribution, where γ is the specified level of significance of the test. Tests against an alternative of decreasing trend, or against a two-sided alternative hypothesis, may be performed in an analogous manner.

Mann's test may be shown to be equivalent to the use of Kendall's rank correlation coefficient to test for the independence of the y_j's with the order in which they are observed (28). Kendall's coefficient is available in SAS PROC CORR. The function BRANKS in SAS PROC MATRIX may be also used to compute either K or Kendall's coefficient.

Mann's test as stated above gives a test for trend to be applied to a fixed set of n observations. In order to apply it for our purposes, it must be adapted for use in a sequential manner. One heuristic method of doing so is outlined below. For definiteness, the procedure will be described as it would be performed in a situation where permeability is presumed to increase to its final equilibrium level. The modifications required in cases of decreasing permeability, or in cases where either increasing or decreasing behavior might occur, are immediate.

1) The permeability test is permitted to run over a preliminary period of at least one pore volume.

2) Following the preliminary period, log-permeabilities y_t are measured at n equally spaced points in time $(t = 1,2,...,n)$.

3) Mann's test is used to test the hypothesis of no trend against the alternative hypothesis of increasing trend at a specified level of significance γ_1. If the hypothesis of no trend is rejected, then the oldest observation is deleted and a new observation is made. Then, step 3) is repeated until the hypothesis of no trend is accepted, at which point step 4) is conducted.

4) Suppose $\beta^* > 0$ represents a bound specified by the experimenter, such that he or she wishes to be sure that the actual slope of the relationship at termination is less than β^*. Adjust the current data by computing $y^*_t = y_t - \beta^* t$. Then use Mann's test on the adjusted data. If the null hypothesis that the adjusted data show no trend is rejected at the level of significance γ_2 against an alternative hypothesis of downward trend, then terminate sampling. Otherwise, make an additional observation, but do not drop the oldest observation, thereby increasing the current sample size by one. Then return to step 3).

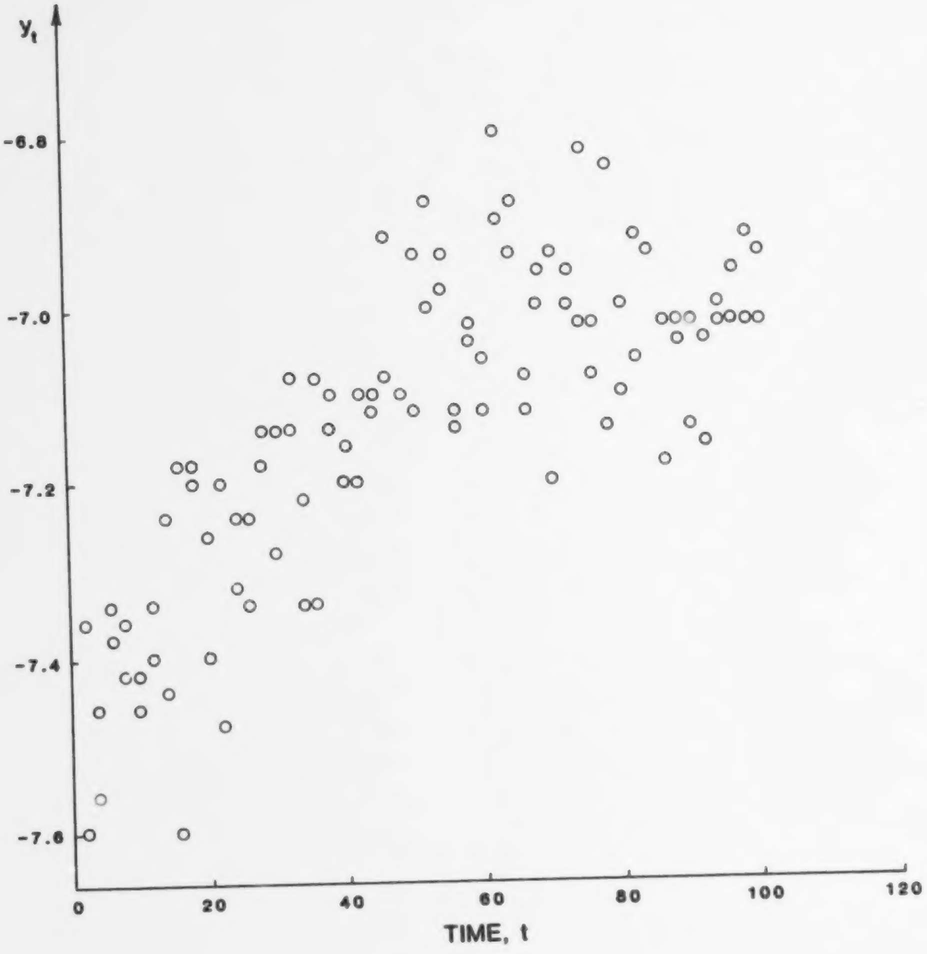
The first three steps of this procedure are extremely similar to Peirce's (41) proposed method, except that Mann's statistic is used in place of linear regression. The rationale for this is that Mann's statistic is designed to be sensitive to any increasing trend, whereas linear regression may be relatively insensitive in a case where, for example, permeability is increasing at a decreasing rate as equilibrium is approached. Mann's statistic is also somewhat less sensitive to occasional unusual observations which might unduly affect a slope computed by least squares, and does not require the assumption of normally distributed within-test errors.

The fourth step of the procedure is included to guard against premature termination which will be caused if a trend is obscured by considerable noise in the observed data. An acceptance of the hypothesis of no trend in step 3) does not necessarily provide strong evidence that no trend exists, but rather only implies that such a hypothesis is not implausible. Thus in step 4) a check is made to see whether the procedure is sensitive enough to detect increasing trends of meaningful magnitude. If so, termination is called for. Otherwise, the procedure adaptively increases sample size to increase the sensitivity of future tests.

The procedure may be illustrated by applying it to the data of Figure C.1. In the figure, the point labeled $t = 0$ corresponds to the end of the preliminary period of step 1). Using an initial sample size of $n = 15$, nominal levels of significance $\gamma_1 = \gamma_2 = .05$ and the bound $\beta^* = 0.004$, the procedure first calls for termination at time $t = 80$. By that time, repeated application of step 4) in the procedure had increased the sample size from 15 to 38. The trend over the interval [43,80] was judged insignificant by Mann's test, and in addition, Mann's test applied to the adjusted data $y_t - .004t$ rejected the hypothesis that the slope of the trend was as large as 0.004. In fact, these data were simulated in such a way that the true point at which equilibration occurs is $\tau = 50$, in order to test the proposed procedure under known conditions.

It is worthwhile to remark that had step 4) been omitted from the algorithm, the use of the step 3) criterion alone (insignificant test for trend) would have caused the procedure to call for termination in 11 of the first 22 iterations of the algorithm, well before the true equilibrium point was reached. This illustrates that with relatively noisy data the insignificance of a test for trend does not provide a reasonable termination criterion, unless coupled with some mechanism which ensures that adequate sensitivity against trends of meaningful magnitude will be achieved.

In order to apply this method, the initial sample size n, the nominal levels of significance γ_1 and γ_2, and the bound β^* must be specified. Limited experience with the procedure indicates that its performance is not sensitive to the choice of n, as would be expected because step 4) adaptively increases the sample size as the algorithm proceeds. We would recommend that n be initially set at least equal to 10. Except for very permeable clays, we recommend that observations be spaced at least one day apart to reduce the effect of serial correlation of observations, which is not explicitly considered in the development given above. Impermeable clays would probably require a longer time interval between observations to permit enough flow to minimize measurement error.



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure C.1 Simulated Test Data Set 1 for the equilibration problem

Because of the sequential (and essentially heuristic) nature of the procedure, it does not appear feasible to determine the actual probability that termination will be prematurely indicated. A conservative procedure (one which reduces the risk of premature termination) is obtained by making the step 3) level of significance γ_1 large and the step 4) level of significance γ_2 small.

It is to be expected that the procedure will be rather sensitive to the step 4) bound β^*. For example, reducing β^* from .004 to .002 in the analysis of the data of Figure 78 delays termination by an additional 11 observations. This is unavoidable, since β^* is a quantification of the degree of trend the experimenter judges to be practically (as opposed to statistically) significant.

In research experiments, comparative permeability tests should, whenever possible, be run concurrently, as discussed in previous Sections. It is in principal possible to base a joint termination criterion on multivariate extensions of Mann's test as developed in (2) and (17). However, at present it seems most advisable to monitor each test individually as described above, but to defer termination of all tests until each has satisfied the termination criterion at least once.

Bayesian Analysis of the Equilibration Problem:

The adaptation of Mann's test described above is essentially a heuristic procedure. A methodology with a somewhat more unified theoretical basis can be developed by addressing the equilibration problem within the framework of Bayesian statistics. Such a framework will be outlined in the remainder of this Section. The methods to be described have certain limitations which might compromise their immediate usefulness, but the authors believe that the potential for refinement of these methods is sufficient to warrant their discussion in order to provide a basis for future research.

Consider once again the problem of analyzing data such as those shown in Figure 78, with the goal of determining the point in time at which equilibration occurs. One method of approaching this question involves postulating a model of the form

$$y_t = \begin{cases} \alpha + f(t-\tau;\ \underset{\sim}{\beta}) + \varepsilon_t, & t = 1,2,\ldots,\tau \\ \alpha + \varepsilon_t & , t \geq \tau \end{cases} \qquad (12)$$

where y_t represents the observation made at time t, τ is the unknown time at which equilibrium is achieved, α is the expected value of y_t after the achievement of equilibrium and $f(\cdot;\beta)$ is a function of time depending on a vector of parameters $\beta = (\beta_1, \beta_2, \ldots, \beta_r)$ which must be estimated from the data. Continuity of the y_t versus t relationship requires that $f(0;\beta) = 0$. ε_t, $t = 1,2,\ldots$, are unobservable error terms.

At any time n, observations $y_1, y_2, \ldots, y_n$ are available. By fitting model (12) to the data, the parameters α, β, and τ may be estimated. This in turn would allow an assessment of the plausibility that $\tau > n$. If the data indicate that this event is extremely unlikely, then sampling may be terminated; otherwise an additional observation is obtained and the process repeated until termination is called for.

This general approach is conveniently implemented within a Bayesian framework. If a prior distribution is specified for the unknown parameters in model (12), including the unknown equilibrium point τ, then after each new observation is obtained, this distribution may be updated by means of Bayes' Theorem. If the loss incurred by inaccurate estimation of the equilibrium level due to premature termination could be quantified, together with the cost of obtaining additional observations, a stopping rule could be developed to minimize the expected total loss. A somewhat simpler alternative would be to adopt a stopping rule which calls for observation to be terminated at time n as soon as the posterior probability that τ exceeds n is less than a specified bound.

If the function f in model (12) is linear in the unknown parameters $\underset{\sim}{\beta}$ then the model may be expressed as

$$y_t = \begin{cases} \alpha + \sum_{j=1}^{r} \beta_j x_j(t-\tau) + \varepsilon_t, \ t = 1,2,\ldots,\tau \\ \\ \alpha + \varepsilon_t, \ t \geq \tau \end{cases} \tag{13}$$

where $x_1(t-\tau), x_2(t-\tau), \ldots, x_r(t-\tau)$ are specified functions of $t-\tau$. For example, in a model in which the approach to equilibrium is linear, r would be 1 and $x_1(t-\tau)=(t-\tau)$, $t = 1,2,\ldots,\tau$. A model permiting curvature in the approach to equilibrium would take $r = 2$ with $x_1(t-\tau)=(t-\tau)$, $x_2(t-\tau) = (t-\tau)^2$.

It will be assumed that the unobservable error terms ε_t in (13) are independently normally distributed with common variance σ_ε^2. Let p_t, $t=r+1, r+2,\ldots$ denote the prior probability function of τ (to avoid certain technical complications in this discussion, the nonrestrictive assumption is made that equilibrium is not reached until at least r+1 observations are made).

Under these conditions, the computation of the posterior distribution of $(\tau,\alpha,\beta_1,\ldots,\beta_r,\sigma_\varepsilon^2)$ becomes rather straightforward if the prior density of $(\alpha,\beta_1,\ldots,\beta_r,\sigma_\varepsilon^2)$, given τ, is a member of the multivariate normal-inverted gamma family. In the development given here, the joint prior of $(\alpha,\beta_1,\ldots,\beta_r,\sigma_\varepsilon^2)$ given τ will be taken to be the (improper) density

$$g(\alpha,\beta_1,\ldots,\beta_r,\sigma_\varepsilon^2|\tau) = \sigma_\varepsilon^{-2}, \ -\infty < \alpha < \infty, \ -\infty < \beta_j < \infty, \ \sigma_\varepsilon^2 > 0$$

in order to model the situation where prior information on these parameters is vague. In this case, Bayes' Theorem may be used to show that the posterior distribution of the equilibrium point τ, given the observation $\underset{\sim}{y}_n = (y_1,y_2,\ldots,y_n)$, is

$$P\{\tau = t \mid \underset{\sim}{y}_n\} = h_t(n) \Big/ \sum_{j=r+1}^{\infty} h_j(n) \qquad , t = r+1, r+2, \ldots$$

where $h_t(n) =$

$$p_t \left(\frac{|C_n|}{|C_t|\left(1+\frac{(n-t)}{w_t}\right)} \right)^{\frac{1}{2}} \left(\frac{S_n}{S_t + \sum_{j=t+1}^{n} (y_j - \bar{y}_{n-t})^2 + (n-t)\left(1+\frac{(n-t)}{w_t}\right)^{-1} (\bar{y}_{n-t} - \hat{\alpha}_t)^2} \right)^{\frac{(n-r-1)}{2}}, \qquad (14)$$

$$r+1 \le t \le n,$$

and $h_t(n) = p_t$, $t \ge n$.

In equation (14), the following notation is used: Let $X(m)$ denote the matrix

$$X(m) = \begin{pmatrix} 1 & x_1(1-m) & x_2(1-m) & \cdots & x_r(1-m) \\ 1 & x_1(2-m) & x_2(2-m) & \cdots & x_r(2-m) \\ 1 & x_1(3-m) & x_2(3-m) & \cdots & x_r(3-m) \\ 1 & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 1 & x_1(0) & x_2(0) & \cdots & x_r(0) \end{pmatrix}.$$

Then, $C_m = X(m)^T X(m)$, $|C_m|$ is the determinant of C_m and w_m is the reciprocal of the upper left-hand element of C_m^{-1}. $\hat{\alpha}_m$ and S_m are respectively the least squares estimate of α and the error sum of squares obtained by fitting the regression model

$$y_t = \alpha + \sum_{j=1}^{r} \beta_j x_j(t-m) + \varepsilon_t, \quad t = 1, 2, \ldots, m$$

to the data $y_1, y_2, \ldots, y_m$ by least squares. $\bar{y}_{n-t}$ represents the average $\sum_{j=t+1}^{n} y_j/(n-t)$.

Equation (14) shows that the posterior distribution of τ may be updated at time $t=n$ by simply fitting a regression model of the form

$$y_t = \alpha + \sum_{j=1}^{r} \beta_j x_j(t-n) + \varepsilon_t$$

to the available data $y_1, y_2, \ldots, y_n$, in order to obtain $\hat{\alpha}_n$ and S_n. The related quantities $|C_n|$ and w_n are by-products of the required computations.

In practice the amount of computation is minimized by recursively fitting these regressions (i.e. obtaining the results of the (n+1)st regression by updating the nth regression) as described in (9), and by the fact that for functions $x_j(t-n)$ of interest (e.g. linear or quadratic trend) it is straightforward to express $|C_n|$ and w_n as explicit functions of n (9).

The general approach outlined above may be illustrated by applying it to the data of Figure 78. A model of the form

$$y_t = \begin{cases} \alpha + \beta(t-\tau) + \varepsilon_t & t = 1, 2, \ldots, \tau \\ \alpha + \varepsilon_t & t \geq \tau \end{cases} \tag{15}$$

was adopted for use in this example, and the prior distribution of τ was taken to be geometric for $\tau \geq 2$ with a mean of 100. For each $n \geq 2$, the posterior probability that $\tau > n$ was computed from equation (14) and

$$P(\tau > n | \underset{\sim}{y}_n) = \sum_{t=n+1}^{\infty} p_t \Bigg/ \left(\sum_{t=2}^{n} h_t(n) + \sum_{t=n+1}^{\infty} p_t \right)$$

(in practice, this would be done as each observation became available). The stopping rule calling for termination of sampling as soon as $P(\tau>n|\underset{\sim}{y}_n)<.01$ was adopted. This condition was satisfied for the first time when n=79. The posterior distribution of τ at that point is shown in Figure C.2. The posterior probability that τ lies in the interval [45, 63] is about 95%, with the highest posterior probability occurring at τ=52. As previously described, these data were simulated, and the true point at which equilibrium was achieved was set at τ=50.

A practical difficulty in the use of such a method lies in the specification of the prior probability distribution of τ. So far, our very limited experience with simulated data indicates that the stopping rule described here may be insensitive to the choice of prior distribution. For example, use of a geometric prior with a mean of 50 in the above example led to termination after 78 observations. A geometric prior with a mean of 20 led to termination after 75 observations.

Unfortunately, the stopping rule is rather sensitive to the functional form which is specified for $f(t-\tau;\underset{\sim}{\beta})$ in equation (12). Shown in Figure C.3 are data which have fairly apparent curvature in the approach to equilibrium. These data were simulated. The point at which equilibrium is achieved is τ=50, and for $t<\tau$ the true relationship is quadratic. If the model

$$y_t = \begin{matrix} \alpha + \beta_1(t-\tau) + \beta_2(t-\tau)^2 + \varepsilon_t, \ t < \tau \\ \alpha + \varepsilon_t, \ t \geq \tau \end{matrix}$$

is used to fit the data, the stopping rule $P\{\tau>n|\underset{\sim}{y}_n\} < .01$ calls for termination after the 83rd observation. The value of τ having highest posterior probability is τ=46, and 95% of the posterior distribution of τ falls in the interval [38, 67]. Thus, as before, the procedure estimates the true equilibrium point rather closely. On the other hand, if the curvature of the data had not been taken into account, and the linear model (15) were used

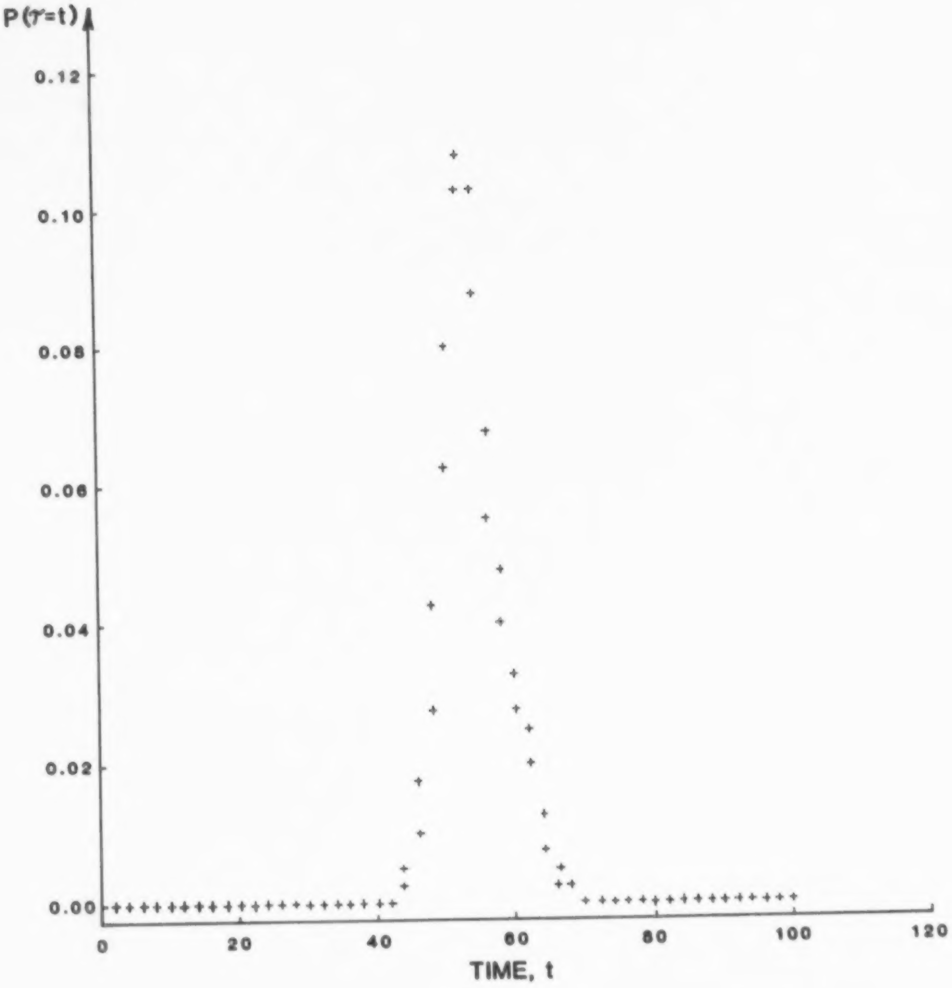
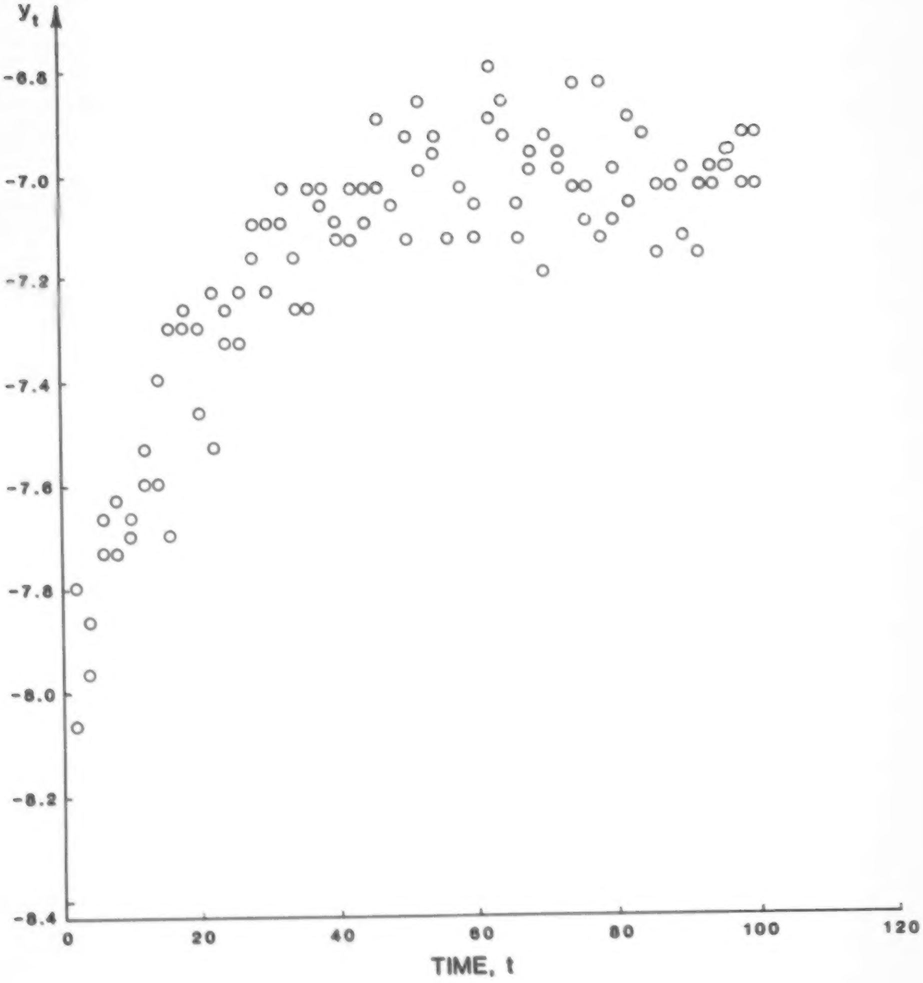


Figure C.2 Posterior probability distribution of the time at which equilibrium is reached (Data Set 1)



(NOTE: SEE APPENDIX A FOR CONVERSION OF LOG_{10} PERMEABILITIES)

Figure C.3 Simulated Test Data Set 2 for the equilibration problem

to generate the posterior probabilities, the stopping rule would have called for termination after 42 observations, falling somewhat short of the actual equilibrium point. Several methods of reducing the effect of such model sensitivity are currently under investigation, including the heuristic device of applying the linear model (15) to the data after having replaced the observed y-values by their ranks.

In the application of this method to actual data from permeability tests, it would be advisable to defer the initiation of the procedure until the tests have been running long enough for a noticable monotone time trend to develop. Thus, as an example, the tests due to Daniel shown in Figure 32 display first decreasing, then increasing, permeability over the first pore volume, after which the pattern is one of monotone increase to the equilibrium level. In such a case one might let the time t=0 represent the point at which one pore volume of flow is achieved.